Effectiveness and Adverse Effects of Deep Brain Stimulation: Umbrella Review of Meta-Analyses

Papageorgiou, Panagiotis N ; Deschner, James ; Papageorgiou, Spyridon N

Abstract: Background This umbrella review summarizes the evidence across meta-analyses regarding the effectiveness and adverse effects of deep brain stimulation (DBS). Methods Databases were searched up to March 2015 for meta-analyses of comparative trials in humans assessing the effectiveness or adverse effects of DBS. Data selection, data extraction, and risk of bias assessment were performed by two independent reviewers. Results Seven eligible systematic reviews were included assessing the use of DBS for epilepsy (n = 1), obsessive-compulsive disorder (n = 1), and Parkinson disease (n = 5). The summary estimates were significant at p 0.05 in four meta-analyses (27%) with both fixed and random effects. One meta-analysis reported that DBS was more effective than sham in reducing the Yale-Brown Obsessive Compulsive Scale score in obsessive-compulsive disorder patients. The remaining three meta-analyses reported differences regarding mortality and depression in patients with Parkinson disease between DBS of the subthalamic nucleus and of the globus pallidus internus. Of the 15 meta-analyses, none compiled adequately robust evidence. Conclusions Although DBS has emerged as a viable surgical intervention to treat various disabling neurologic symptoms, existing studies fail to adequately support its use based on robust evidence without hints of bias.

DOI: https://doi.org/10.1055/s-0036-1592158

Originally published at:
Papageorgiou, Panagiotis N; Deschner, James; Papageorgiou, Spyridon N (2016). Effectiveness and Adverse Effects of Deep Brain Stimulation: Umbrella Review of Meta-Analyses. Journal of Neurological Surgery. Part A: Central European Neurosurgery, 78(02):180-190.
DOI: https://doi.org/10.1055/s-0036-1592158

Effectiveness and adverse effects of deep brain stimulation: umbrella review of meta-analyses

Panagiotis N. Papageorgiou, MD[1] • James Deschner, DMD, PhD[2] • Spyridon N. Papageorgiou, DDS, Dr med dent[3,4]

[1]Department of Neurosciences, Southampton University Hospital, Tremona Road, SO16 6YD, Southampton, UK
[2]Section of Experimental Dento-Maxillo-Facial Medicine, School of Dentistry, University of Bonn, 53111, Bonn, Germany
[3]Department of Orthodontics, School of Dentistry, University of Bonn, 53111, Bonn, Germany
[4]Department of Oral Technology, School of Dentistry, University of Bonn, 53111, Bonn, Germany

*Corresponding author: Panagiotis N. Papageorgiou, MD, Department of Neurosciences, Southampton University Hospital, Tremona Road, SO16 6YD, Southampton, UK; E-mail: pnpapag@gmail.com.

Words in abstract: 196 / words in text: 3203
Running title: Deep brain stimulation: umbrella review
Conflicts of interest: None.

Authors' contributors: PNP and SNP conceived the study. PNP and SNP searched for and selected the systematic reviews, did the data abstraction and evaluated the included systematic reviews. JD resolved conflicts at any stage. SNP did the statistical analysis. All authors contributed to writing and reviewing the manuscript. PNP is the guarantor.

Keywords: deep brain stimulation; Parkinson's disease; epilepsy; obsessive compulsive disorder; systematic review; meta-analysis

Effectiveness and adverse effects of deep brain stimulation: umbrella review of meta-analyses

Abstract

Background This umbrella review summarizes the evidence across meta-analyses regarding the effectiveness and adverse effects of deep brain stimulation (DBS).

Methods Databases were searched up to March 2015 for meta-analyses of comparative trials in humans assessing the effectiveness or adverse effects of DBS. Data selection, data extraction, and risk of bias assessment was performed by two independent reviewers.

Results Seven eligible systematic reviews were included assessing the use of DBS for epilepsy (n = 1), obsessive-compulsive disorder (n = 1) and Parkinson's disease (n = 5). The summary estimates were significant at $p \leq 0.05$ in 4 meta-analyses (27%) with both fixed- and random-effects. One meta-analysis reported that DBS was more effective than sham in reducing the Yale–Brown Obsessive Compulsive Scale score in obsessive compulsive disorder patients. The remaining three meta-analyses reported differences regarding mortality and depression in patients with Parkinson's disease between DBS of the subthalamic nucleus and of the globus pallidus internus. Of the 15 meta-analyses, none compiled adequately robust evidence.

Conclusions Though DBS has emerged as a viable surgical intervention to treat various disabling neurological symptoms, existing studies fail to adequately support its use based on robust evidence without hints of bias.

Manuscript Text

Introduction

Rationale

Deep brain stimulation (DBS) has emerged as a novel neurosurgical method to treat movement disorders, such as Parkinson's disease,[1] essential tremor,[2-4] chronic pain,[5] Tourette's syndrome,[6] and psychiatric disorders, like obsessive compulsive disorder,[7] depression,[8-10] and anorexia nervosa.[11] High-frequency stimulating electrodes are placed in one of several target areas in the brain, including the ventrolateral thalamus, subthalamic nucleus (STN) or internal segment of the globus pallidus (GPi) and periaqueductal grey matter, and is connected to an implantable pulse generator. DBS involves the delivery of precise electrical signals to specific deep anatomical structures in the central nervous system. DBS is thought to affect the firing rates and bursting patterns of neurons and, ultimately, the synchronized oscillatory activity of neuronal networks.[12-13]

Objectives

For some of the abovementioned neurological conditions, several randomized controlled trials and systematic reviews thereof have been published in the last decade, but have not been systematically evaluated up to now. To evaluate the strength of evidence regarding the clinical indication for the use of DBS, we performed an umbrella review of the evidence across meta-analyses pertaining to the effectiveness or adverse effects of DBS. We aimed to assess the direction and magnitude of existing effects, as well as evaluate whether there are hints of biases in the literature that could endanger the validity of the results.

Materials and Methods

Literature search

Two researchers (PNP and SNP) independently searched MEDLINe through PubMed, Scopus, the Cochrane Database of Systematic Reviews, and the Cochrane Database of Abstracts of Reviews of Effects from inception to the end of March 2015 for meta-analyses or systematic reviews of studies investigating the effectiveness or adverse effects of deep brain stimulation. The exact literature search for each database can be seen in Appendix A. The references from eligible systematic or narrative reviews were also checked. The titles, abstracts, and full-texts of the resulting papers were examined in detail, and discrepancies were resolved by a third researcher (JD).

Eligibility criteria and data extraction

Articles were eligible, if the authors had performed a systematic search to identify pertinent clinical trials on DBS in humans and had performed quantitative data synthesis comparing at least two experimental/control groups. We included systematic reviews on both randomised controlled trials (RCT) and non-RCTs. Meta-analyses or systematic reviews that did not present study specific data were excluded, but the data were requested from corresponding author. We included meta-analyses of both binary and continuous outcomes. If an article presented separate meta-analyses on more than one eligible outcome, those were assessed separately. Whenever more than one meta-analysis existed on the same scientific question, the meta-analysis with the largest number of

studies and/or the most complete reporting was selected, but we conducted sensitivity analyses to assess any differences in these duplicate meta-analyses.

From each eligible systematic review, two authors (PNP and SNP) abstracted independently information on publication type, number of searched databases, search period, type of included trials, number of assessed outcomes (including measures against multiple testing), and citation counts from Google Scholar (http://scholar.google.com). We recorded if the included systematic reviews assessed the risk of bias of the individual studies and the quality of evidence according to the GRADE approach,[14] but we did not perform these procedures ourselves, as this task was beyond the scope of this umbrella review. Additionally, we appraised the methodological quality of the included systematic reviews with the AMSTAR tool.[15] Finally, we assessed the risk of bias of the included systematic reviews with the newly-designed ROBIS tool.[16]

From each eligible meta-analysis, the same two authors abstracted information independently on first author, year of publication, outcome examined, number of included studies, and reported data at the individual trial level. For each of the included studies in each eligible meta-analysis, we recorded the study design (RCT or non-RCT), the number of cases (for binary outcomes) and population participants.

Assessment of summary effects and heterogeneity

For meta-analyses of continuous outcomes, Standardized Mean Differences (SMDs) were chosen as effect estimates. Binary outcomes were also transformed to SMDs in order to enable synthesis of both continuous and binary outcomes together.

We estimated the summary effects using both fixed-effect and inverse variance random-effects models.[17] Fixed-effect meta-analysis relies on the assumption that a unique effect underlies every study in the meta-analysis and no heterogeneity between studies exists. A random-effects synthesis makes the assumption that individual studies are estimating different effects, which are assumed to have a normal distribution. The random-effects meta-analysis is performed to estimate the mean of this distribution of effects across different studies and the uncertainty about that mean (95% confidence interval (CI)). We also calculated the 95% prediction interval (PrI) for the summary random-effects estimates, which further account for heterogeneity between studies and indicate the uncertainty for the effect that would be expected in a new study examining that same association.[18] The 95% PrI shows where the true effects are for 95% of the studies from the population of studies that are synthesised or similar (exchangeable) studies that might be done in the future.

We assessed heterogeneity between studies using the P value of the χ^2 based Cochran Q test and the I^2 metric of inconsistency; this could reflect either genuine diversity or bias. The Q test is obtained by the weighted sum of the squared differences of the observed effect in each study minus the fixed summary effect.[19] The I^2 metric ranges between 0% and 100% and is the ratio of variance between studies over the sum of the variances within and between studies. The 95 % CIs around I^2 were calculated according to the non-central chi-square approximation of Q.[20]

A sensitivity analysis according to the basic study design of included trials[21] was conducted with mixed-effects subgroup analysis (random-effects meta-regression) and by calculating the ΔSMDs (difference in SMDs) and the associated 95% CIs. An iterative

residual maximum likelihood algorithm was used for the estimation of between-study variance because of its performance,[22] while the Knapp-Hartung modification[23] was used for the calculation of the ΔSMDs, which accounts for the uncertainty in the heterogeneity estimate.[24] The effect magnitude both for SMD and ΔSMD was conventionally judged as 0.2=small effect, 0.5=medium effect, and 0.8= large effect.[25] The cut-off of SMD or ΔSMD>0.8 was used to construct contours of large effect magnitude in all forest plots.

Assessment of small study effects

We examined whether there is an indication for small study effects - that is, if small studies tend to give higher estimates than large studies. Small study effects can indicate publication bias or other reporting biases, but they can also reflect genuine heterogeneity, chance, or other reasons for differences between small and large studies.[26] We used the regression asymmetry test proposed by Egger[27] to investigate funnel plot asymmetry. The alternative test proposed by Harbord outperforms Egger test, but is applicable only to dichotomous outcomes.[28]

Associations meeting further criteria

We further identified associations for which the summary fixed and random-effects estimates showed strong evidence of significance ($P<0.001$; a threshold that has been suggested to substantially reduce the number of false positive findings[29]) were based on evidence from more than 500 cases for binary outcomes or from more than 5000 participants for continuous outcomes;[30] did not have large heterogeneity between studies ($I^2<75\%$); their 95% PrI excluded the null value; and had no evidence of small study effects

(from Egger test). All calculations were performed in STATA version 12. All P values are two-sided, while statistical significance is set at 5% for all tests, except for heterogeneity and Egger tests (10%). Although multiple P values are reported in this umbrella review, no correction or adjustment was performed, as we aimed to summarize the results of the included papers.

Results

Characteristics of included systematic reviews

A total of 100 hits were received from the electronic literature search, while another 3 papers were added manually (Fig. 1). After removal of duplicates, screening and application of the inclusion criteria, a total of 7 systematic reviews were included in this umbrella review[31-37] (Table 1; Appendix B). A total of 18 reviews with potentially eligible meta-analyses were excluded from the umbrella review, as they did not include conventional pairwise meta-analytic comparisons, but performed averaging among studies, pooled single group estimates, or generated control group data from expert opinions. Missing data at the trial level were requested in two instances (Appendix B), but no answer was received and the reviews were excluded. Chambers and Bowen[32] performed meta-analysis of both DBS and vagus nerve stimulation at various stimulation degrees for the treatment of epilepsy, but only the subgroup on DBS (one single trial) is included in this umbrella review.

Four reviews were published in scientific journals[32-34,36] one was published in the Cochrane Library,[37] while the rest two were posters from a scientific congress (Appendix

B).[31,35] All reviews were in English, published after 2013 and searched 1-4 literature databases. Five reviews[32-34,36-37] included only RCTs, while remaining two[31,35] included both RCTs and non-RCTs. Four reviews[31,34,36,37] reported more than one primary outcomes in their analyses, but no measure was taken to safeguard against increased Type I errors.

Risk of bias and methodological adequacy

Out of the seven included systematic reviews, 4 of them[32-34,37] (57%) assessed the risk of bias in the included trials; all of them with the Cochrane Collaboration's risk of bias tool. Most often sources of bias identified concerned the generation of random sequence,[33] allocation concealment,[34] blinding,[34,37] incomplete outcome data,[37] selective reporting,[32,33,37] and potentional confounding due to uncontrolled simultaneous drug administration.[37]

The quality of evidence (strength of recommendations) from the performed meta-analyses was assessed with the GRADE approach only in 2 out of 7 reviews[32,37] (29%). The strength of recommendations from the one review comparing on vs off DBS for drug-resistant epilepsy was low, due to imprecision and hints of publication bias.[32] The strength of recommendations from the review comparing DBS to sham ranged from very low to high, with imprecision, confounding due to missing wash-out period in cross-over trials and absence of drug regulation being the reasons to downgrade the quality.[37]

The methodological adequacy of the included reviews was appraised with the AMSTAR tool. This appraisal was limited only to reviews published in journals or in the Cochrane Library, as the two reviews in poster format did not provide adequate information. The AMSTAR scores for the five reviews ranged from 2 to 11 (Appendix C-D),

with only the Cochrane Review scoring the maximum of 11 points. The main shortcomings were lack of a priori design, incomplete reporting of included/excluded studies, and missing statements for possible conflicts of interest.

The risk of bias of the included reviews was appraised with the ROBIS tool. As with the appraisal of methodological adequacy, risk of bias was assessed only in reviews published in journals or in the Cochrane Library. Four out of five reviews were judged to be in high risk of bias (Fig. 2; Table 2; Appendix E), with only the Cochrane Review being in low risk of bias.

Summary effect sizes

A total of 15 unique meta-analyses were extracted from the seven included review: four meta-analyses (from one review[37]) regarding DBS use for drug-resistant epilepsy, one meta-analysis (from one review[33]) regarding DBS use for obsessive compulsive disorder, and ten meta-analyses (from 5 reviews[31,32,34-36]) regarding DBS use for Parkinson's disease. For clarity reasons the results of the included meta-analyses are presented separately concerning the efficacy (comparison of active treatment with sham treatment; Table 3; Fig. 3) and comparative effectiveness (comparison of two active treatments; Table 4; Fig. 4) of DBS.

As far as the efficacy of DBS is concerned the results of the meta-analyses were significant at the $P \leq 0.05$ level in one instance, both with fixed- and random-effects models: absolute effectiveness of DBS compared to sham in patients with obsessive compulsive disorder. The results of the largest study, compared to the pooled effect, were almost the same in all cases (meta-analyses 3, 8, 10). The 95% PrI did not exclude the null value in any

case, indicating that no clinical recommendations can be done on the basis of existing evidence.

As far as the comparative effectiveness of DBS is concerned the results of the meta-analyses were significant at the P≤0.05 level in three instances, both with fixed- and random-effects models, regarding the effectiveness against depressive symptoms or minimization of mortality with DBS use for Parkinson's disease according to the stimulated target. The results of the meta-analyses were significant at the P<0.001 level with both models regarding the non-motor symptoms of Parkinson's disease, assessed with the Beck Depression Inventory-II. Table 4 also shows the effect of the largest study included in each meta-analysis. The results of the largest study, compared to the pooled effect, agreed in direction but were more conservative in 4 cases (meta-analyses 1, 2, 3, 5), were almost the same in 3 cases (meta-analyses 4, 6, 7) and were on the opposite direction in 3 cases (meta-analyses 8, 9). The 95% PrI excluded the null value only for the meta-analysis of post-DBS depression severity compared to the DBS target (meta-analysis 3).

Overlap among meta-analyses

Overlaps among meta-analyses existed only regarding the use of DBS in patients with Parkinson's disease. Overlaps existed among three reviews for three studied outcomes, i.e. the unified Parkinson's disease rating scale (UPDRS)-II, UPDRS-III off-medication, and UPDRS-III on-medication. The review of Liu et al[34] was selected over the review of Sako et al[36] and Arnaout et al,[31] based on number of included studies and completeness of the analyses, respectively. Both the matrix of included studies and the results of the meta-analyses with overlap can be seen in the Supplement (Appendix F-G).

Heterogeneity and small study effects

The Q test showed significant heterogeneity (P≤0.10) for two meta-analyses (Table 4), both with moderate to high inconsistency (I^2 between 50% and 75%). Another meta-analysis had moderate inconsistency (52%), while the rest had I^2<50%. Uncertainty around the heterogeneity estimates was often large, as reflected by wide 95% CIs of the I^2. In one of the meta-analyses with significant heterogeneity (Table 4; meta-analysis 8) the authors conducted a sensitivity analysis by removing the trial with the largest effect, thereby reducing the heterogeneity. However, no clinical reasoning was provided for this choice. Sensitivity analyses according to the study design of the included trials (RCT or non-RCT) were possible in four meta-analyses from two systematic reviews[31,35] that included both RCTs and non-RCTs (Appendix H). In three out of four meta-analyses, non-RCTs gave hints of effect overestimation, which was however non-significant in all cases (P>0.05). However, the differences in SMDs ranged between 0.07 and 0.74, the latter approaching the cut-off for large effect magnitude (0.8).

There was no evidence of small study effects, except for one case, where the Egger's test was significant (Table 4; P=0.041). In this meta-analysis the change in semantic verbal fluency in Parkinson patients after STN or GPi DBS was more conservative in the larger studies.

Associations meeting further criteria

Of the 15 included meta-analyses, none fulfilled all criteria set (Appendix I) with the greatest problem being the limited number of studies/patients included.

Discussion

Principal findings and possible explanations

We performed an umbrella review to examine the existing evidence from meta-analyses on the use of DBS in patients with various disabling neurological symptoms. Although 15 meta-analyses were finally included, most reported on nominally non-significant effects, both pertaining to the effectiveness of DBS and to the minimization of adverse effects. Only a small minority of meta-analyses (n=4) reported on statistically significant effects, but none of these was based on robust epidemiological evidence.

DBS for obsessive compulsive disorder

Deep brain stimulation was reported to be more effective in alleviating the disease symptoms of patient's with obsessive compulsive disorder (as measured by the Yale–Brown Obsessive Compulsive Scale) compared to sham stimulation.[33] Although the effects were consistent across all five included studies and the random-effects pooled summary indicated a large effect magnitude, clear indications of imprecision existed, which resulted in uncertainty in the estimates. This can be however remedied by the inclusion of additional RCT on this subject that might strengthen the quality of clinical recommendations for the use of DBS in obsessive compulsive patients. It has been reported that, DBS targeted at the nucleus accumbens of patients with obsessive compulsive disorder appears to release dopamine in the striatum, which is associated with increased plasma levels of homovanillic acid and improved clinical symptoms. This suggests that DBS

may compensate for a defective dopaminergic system. These changes hint at a causal role of dopamine in the therapeutic efficacy of DBS, and agree with previous insight into the role of dopamine in obsessive compulsive disorder,[38,39] but further research is needed to confirm this.[40] Finally, as suggested by the American Society for Stereotactic and Functional Neurosurgery, it is imperative that electrophysiological, morphological (neuroimaging), functional, or clinical predictive factors are developed that will enable to identify the most apt candidate patients for this procedure.[41]

DBS for Parkinson's disease

One of the most interesting findings of this umbrella review is that no meta-analysis on the efficacy of DBS for Parkinson's disease was identified and included. As far as comparative effectiveness is concerned, recent studies have reported that there is no significant difference between STN and GPi DBS in improving motor control in patients with Parkinson's disease,[42,43] which is the main outcome of interest in Parkinson's disease. However, existing evidence indicates that significant differences exist between STN or GPi DBS in patients with Parkinson's disease, as far as mortality or depression are concerned.[34-36] STN DBS was associated with higher patient mortality than GPi DBS ($P<0.01$). However, the authors of the meta-analysis reported that since most deaths were due to post-operative complications not directly related to DBS, confounding cannot be excluded.[35] Additionally, it must also be noted, that non-RCTs were also included in this meta-analysis, which might have introduced further bias,[21] while clear signs of imprecision were also present. Furthermore, STN DBS was also associated with greater incidence of depression and with more severe depression, as measured with the Beck Depression Inventory II,

compared to GPi DBS.[34,36] Thus, GPi DBS may be beneficial to treat severe non-motor mood symptoms than STN DBS. Some of these non-motor symptoms can even predate the motor problems[44-46] and are often more disabling and resistant to treatment than motor symptoms, being key determinants of quality of life. However, due to the uncertainty around the existing results that originate from imprecision, no clear distinction between the pros and cons of STN and GPi DBS can be made, and further studies are needed in order to form robust clinical recommendations.

Strengths, weaknesses and future research

Several limitations and caveats should be considered in the interpretation of our findings. Firstly, we included only studies used in certain published meta-analyses and thus might have missed some individual studies, if those were not identified in the original systematic searches. Secondarily, data at the trial level were missing for some meta-analyses and were therefore excluded. Thirdly, most of the included systematic reviews presented serious methodological inadequacies,[47] like inadequate justification for the statistical model[48] and metric used,[49] incomplete assessement of clinical and statistical heterogeneity[47] and partial risk of bias assessment at the individual trial or meta-analysis level. Finally, funnel plot asymmetry was consistently investigated with Egger's test for all meta-analyses, although less than ten trials were included in every case,[26] according to previous umbrella reviews.[30,50,51]

Future research should be directed into planning and conducting blinded RCTs with a priori sample size calculations that will enable an adequately-powered assessment of the benefits and harms of DBS for its various neurological indications. Furthermore, there is a

great number of published RCTs on the use of DBS for various conditions that has not yet been systematically appraised, including refractory partial seizures,[52] medication-refractory cervical dystonia,[53] essential tremor,[54] Alzheimer's disease,[55] and treatment-resistant depression.[56] These could be the focus of future systematic reviews.

Conclusions

In conclusion, although use of DBS for the treatment of various disabling neurological symptoms has been studied to some effect, firm universal conclusions about either its efficacy or comparative effectiveness cannot be drawn. There is limited evidence that DBS is effective in reducing symptoms in obsessive compulsive disorder patients and that GPi DBS is more beneficial than STN DBS in reducing non-motor symptoms in Parkinson's patients. However, these associations do not meet the epidemiological criteria for credibility that were set for this umbrella review and further well-designed studies are needed.

References

1. Limousin P, Krack P, Pollak P, et al. Electrical stimulation of the subthalamic nucleus in advanced Parkinson's disease. N Engl J Med 1998;339(16):1105–1111
2. Kleiner-Fisman G, Herzog J, Fisman DN, et al. Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. Mov Disord 2006;21 Suppl 14:S290–S304
3. Perlmutter JS, Mink JW. Deep brain stimulation. Annu Rev Neurosci 2006;29:229–257
4. Plaha P, Patel NK, Gill SS. Stimulation of the subthalamic region for essential tremor. J Neurosurg 2004;101(1):48–54
5. Bittar RG, Kar-Purkayastha I, Owen SL, et al. Deep brain stimulation for pain relief: a meta-analysis. J Clin Neurosci 2005;12(5):515–519
6. Visser-Vandewalle V, Temel Y, Boon P, et al. Chronic bilateral thalamic stimulation: a new therapeutic approach in intractable Tourette syndrome. Report of three cases. J Neurosurg 2003;99(6):1094–1100
7. Greenberg BD, Gabriels LA, Malone DA Jr, et al. Deep brain stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: worldwide experience. Mol Psychiatry 2010;15(1):64–79
8. Lozano AM, Mayberg HS, Giacobbe P, et al. Subcallosal cingulate gyrus deep brain stimulation for treatment-resistant depression. Biol Psychiatry 2008;64(6):461–467
9. Schlaepfer TE, Bewernick BH. Deep brain stimulation for major depression. Handb Clin Neurol 2013;116:235–243
10. Schlaepfer TE, Lieb K. Deep brain stimulation for treatment of refractory depression. Lancet 2005;366(9495):1420–1422

11. Wu H, Van Dyck-Lippens PJ, Santegoeds R, et al. Deep-brain stimulation for anorexia nervosa. World Neurosurg 2013;80(3-4):S29.e1–S29.e10
12. Hanajima R, Ashby P, Lozano AM, et al. Single pulse stimulation of the human subthalamic nucleus facilitates the motor cortex at short intervals. J Neurophysiol 2004;92(3):1937–1943
13. Uc EY, Follett KA. Deep brain stimulation in movement disorders. Semin Neurol 2007;27(2):170–182
14. Guyatt GH, Oxman AD, Schünemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64(4):380–382
15. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10
16. Whiting P, Savović J, Higgins JPT, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2015; doi:10.1016/j.jclinepi.2015.06.005
17. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177–188
18. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;342:d549
19. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21(11):1539–1558
20. Orsini N, Bottai M, Higgins J, et al. Heterogi: Stata module to quantify heterogeneity in a meta-analysis. Statistical Software Components 2006;

http://www.EconPapers.repec.org/RePEc:boc:bocode:s449201. Accessed 15 March 15 2015

21. Papageorgiou SN, Xavier GM, Cobourne MT. Basic study design influences the results of orthodontic clinical investigations. J Clin Epidemiol 2015;68(12):1512–1522
22. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999;18(20):2693–2708
23. Knapp G, Hartung J. Improved tests for a random-effects meta-regression with a single covariate. Stat Med 2003;22(17):2693–2710
24. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004;23(11):1663–1682
25. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New York: Academic Press; 1988.
26. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomized controlled trials. BMJ 2011;343:d4002
27. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315(7109):629–634
28. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Stat Med 2006;25(20):3443–3457
29. Johnson VE. Revised standards for statistical evidence. Proc Natl Acad Sci U S A 2013;110(48):19313–19317

30. Theodoratou E, Tzoulaki I, Zgaga L, et al. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. BMJ 2014;348:g2035
31. Arnaout M, Negida A, El Ashal G, et al. Meta-analysis comparing Subthalamic and Pallidal deep brain stimulation for patients with Parkinson's disease. Paper presented at: ASiT conference; Feb 27 – Mar 1, 2015; Glasgow, UK
32. Chambers A, Bowen JM. Electrical stimulation for drug-resistant epilepsy: an evidence-based analysis. Ont Health Technol Assess Ser 2013;13(18):1–37
33. Kisely S, Hall K, Siskind D, Frater J, et al. Deep brain stimulation for obsessive-compulsive disorder: A systematic review and meta-analysis. Psychol Med 2014;44(16):3533–3542
34. Liu Y, Li W, Tan C, et al. Meta-analysis comparing deep brain stimulation of the globus pallidus and subthalamic nucleus to treat advanced Parkinson disease. J Neurosurg 2014;121(3):709–718
35. Negida A, Arnaout M, El Ashal G, et al. Meta-analysis of mortality following Subthalamic and Pallidal deep brain stimulation for patients with Parkinson's disease. Paper presented at: ASiT conference; Feb 27 – Mar 1, 2015; Glasgow, UK
36. Sako W, Miyazaki Y, Izumi Y, et al. Which target is best for patients with Parkinson's disease? A meta-analysis of pallidal and subthalamic stimulation. J Neurol Neurosurg Psychiatry 2014;85(9):982–986
37. Sprengers M, Vonck K, Carrette E, et al. Deep brain and cortical stimulation for epilepsy. Cochrane Database Syst Rev 2014;6:CD008497

38. Borcherding BG, Keysor CS, Rapoport JL, et al. Motor/ vocal tics and compulsive behaviors on stimulant drugs: Is there a common vulnerability? Psychiatry Res 1990;33(1):83–94
39. Vulink NC, Denys D, Fluitman SB, et al. Quetiapine augments the effect of citalopram in non-refractory obsessive-compulsive disorder: A randomized, double-blind, placebo-controlled study of 76 patients. J Clin Psychiatr 2009;70(7):1001–1008
40. Figee M, de Koning P, Klaassen S, et al. Deep brain stimulation induces striatal dopamine release in obsessive-compulsive disorder. Biol Psychiatry 2014;75(8):647–652
41. Hamani C, Pilitsis J, Rughani AI, et al. Deep brain stimulation for obsessive-compulsive disorder: systematic review and evidence-based guideline sponsored by the American Society for Stereotactic and Functional Neurosurgery and the Congress of Neurological Surgeons (CNS) and endorsed by the CNS and American Association of Neurological Surgeons. Neurosurgery 2014;75(4):327–333
42. Follett KA, Weaver FM, Stern M, et al. Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. N Engl J Med 2010;362(22):2077–2091
43. Odekerken VJ, van Laar T, Staal MJ, et al. Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. Lancet Neurol 2013;12(1):37–44
44. Doorn KJ, Lucassen PJ, Boddeke HW, et al. Emerging roles of microglial activation and non-motor symptoms in Parkinson's disease. Prog Neurobiol 2012;98(2):222–238
45. Ferrer I, López-Gonzalez I, Carmona M, et al. Neurochemistry and the non-motor aspects of PD. Neurobiol Dis 2012;46(3):508–526

46. Meissner WG. When does Parkinson's disease begin? From prodromal disease to motor signs. Rev Neurol (Paris) 2012;168(11):809–814
47. Papageorgiou SN. Meta-analysis for orthodontists: Part II--Is all that glitters gold? J Orthod 2014;41(4):327–336
48. Papageorgiou SN. Meta-analysis for orthodontists: Part I--How to choose effect measure and statistical model. J Orthod 2014;41(4):317–326
49. Papageorgiou SN, Tsiranidou E, Antonoglou GN, et al. Choice of effect measure for meta-analyses of dichotomous outcomes influenced the identified heterogeneity and direction of small-study effects. J Clin Epidemiol 2015;68(5):534–541
50. Belbasis L, Bellou V, Evangelou E, et al. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. Lancet Neurol 2015;14(3):263–273
51. Tsilidis KK, Kasimis JC, Lopez DS, et al. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. BMJ 2015;350:g7607.
52. Bergey GK, Morrell MJ, Mizrahi EM, et al. Long-term treatment with responsive brain stimulation in adults with refractory partial seizures. Neurology 2015;84(8):810–817
53. Volkmann J, Mueller J, Deuschl G, et al. Pallidal neurostimulation in patients with medication-refractory cervical dystonia: a randomised, sham-controlled trial. Lancet Neurol 2014;13(9):875–884
54. Pedrosa DJ, Auth M, Pauls KA, et al. Verbal fluency in essential tremor patients: the effects of deep brain stimulation. Brain Stimul 2014;7(3):359–364
55. Olazarán J, González B, López-Álvarez J, et al. Motor effects of REAC in advanced Alzheimer's disease: results from a pilot trial. J Alzheimers Dis 2013;36(2):297–302

56. Ramasubbu R, Anderson S, Haffenden A, et al. Double-blind optimization of subcallosal cingulate deep brain stimulation for treatment-resistant depression: a pilot study. J Psychiatry Neurosci 2013;38(5):325–332

Figure Legends

Fig. 1 Flowdiagram for the study selection. Number of the systematic reviews identified from the electronic search and assessed in this study, together with the number of meta-analyses extracted from them.

Fig. 2 ROBIS risk of bias of the included systematic reviews. Assessment of the risk of bias regarding the four domains covered by the ROBIS tool: the study eligibility criteria, the identification and selection of studies, the data collection and study appraisal, and the synthesis of findings. Darker colours indicate overall risk of bias rating; lighter colors concern judgments.

Fig. 3 Forest plot of all included meta-analyses of deep brain stimulation controlled trials on efficacy with standardized mean difference as type of metric. Summary of the result of all meta-analyses replicated from the included systematic reviews. Abbreviations: SMD, standardized mean difference; DBS, deep brain stimulation; YBOCS, Yale–Brown Obsessive Compulsive Scale.

Fig. 4 Forest plot of all included meta-analyses of deep brain stimulation controlled trials on comparative effectiveness with standardized mean difference as type of metric. Summary of the result of all meta-analyses replicated from the included systematic reviews. Abbreviations: SMD, standardized mean difference; DBS, deep brain stimulation; STN, subthalamic nucleus; GPi, globus pallidus internus; BDI-II, Beck depression inventory-II; LED, levodopa equivalent dose; UPDRS, unified Parkinson's disease rating scale.

Table Legends

Table 1 Characteristics of the included systematic reviews**.** Abbreviations: RCT, randomized controlled trial.

Table 2 Tabular presentation for the ROBIS results of the included systematic reviews. Abbreviations: ⊕, low risk; ⊖, high risk; ?, unclear risk.

Table 3 Results of the included meta-analyses on efficacy (comparison of active interventions with sham controls). * $p<0.05$; ** $p<0.01$; *** $p<0.001$; ¥ significance results originating from the test for Q. Abbreviations: FE, fixed-effect model; RE, random-effects model; N, number of studies; SMD, standardized mean difference; CI, confidence interval; PrI, predictive interval; DBS, deep brain stimulation; Bin, binary; Effect, effectiveness.

Table 4 Results of the included meta-analyses on comparative effectiveness (comparison of two active interventions). * $p<0.05$; ** $p<0.01$; *** $p<0.001$; †there was overlap with the poster report of Arnaout 2015, but the journal report was retained, due to completeness. ‡there was overlap with the poster report of Arnaout 2015, and the journal report of Sako 2014, but the journal report of Liu 2014, was retained, due to completeness and increased number of included studies. ¥ significance results originating from the test for Q. Abbreviations: FE, fixed-effect model; RE, random-effects model; N, number of studies; SMD, standardized mean difference; CI, confidence interval; PrI, predictive interval; DBS, deep brain stimulation; Bin, binary; Effect, effectiveness; NA, not applicable; STN, subthalamic nucleus; GPi, globus pallidus internus; AE, adverse effects; BDI-II, Beck depression inventory-

II; LED, levodopa equivalent dose; UPDRS, unified Parkinson's disease rating scale.

Fig. 1 Flowdiagram for the study selection. Number of the systematic reviews identified from the electronic search and assessed in this study, together with the number of meta-analyses extracted from them.

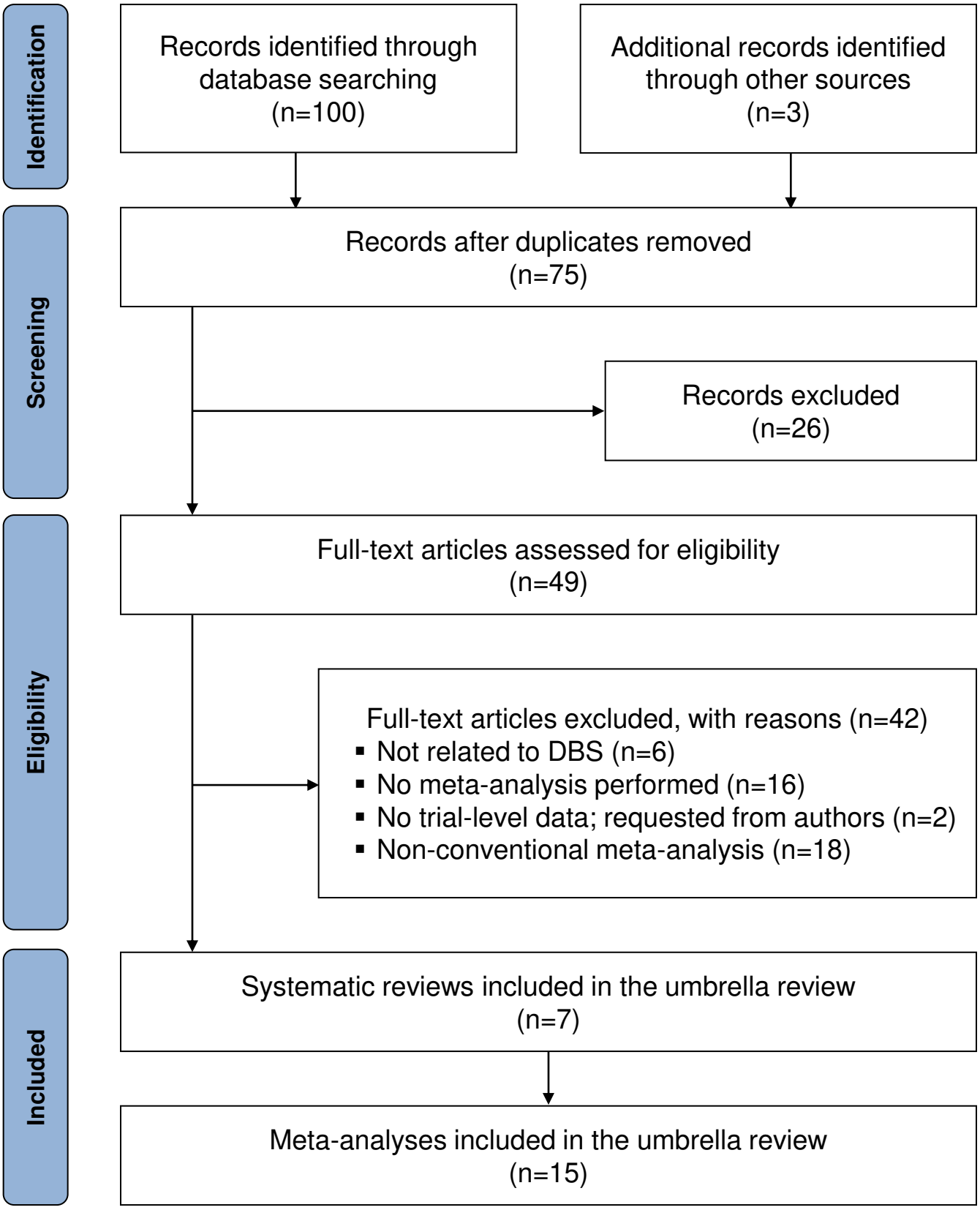


Fig. 2 ROBIS risk of bias of the included systematic reviews. Assessment of the risk of bias regarding the four domains covered by the ROBIS tool: the study eligibility criteria, the identification and selection of studies, the data collection and study appraisal, and the synthesis of findings. Darker colours indicate overall risk of bias rating; lighter colors concern judgments.

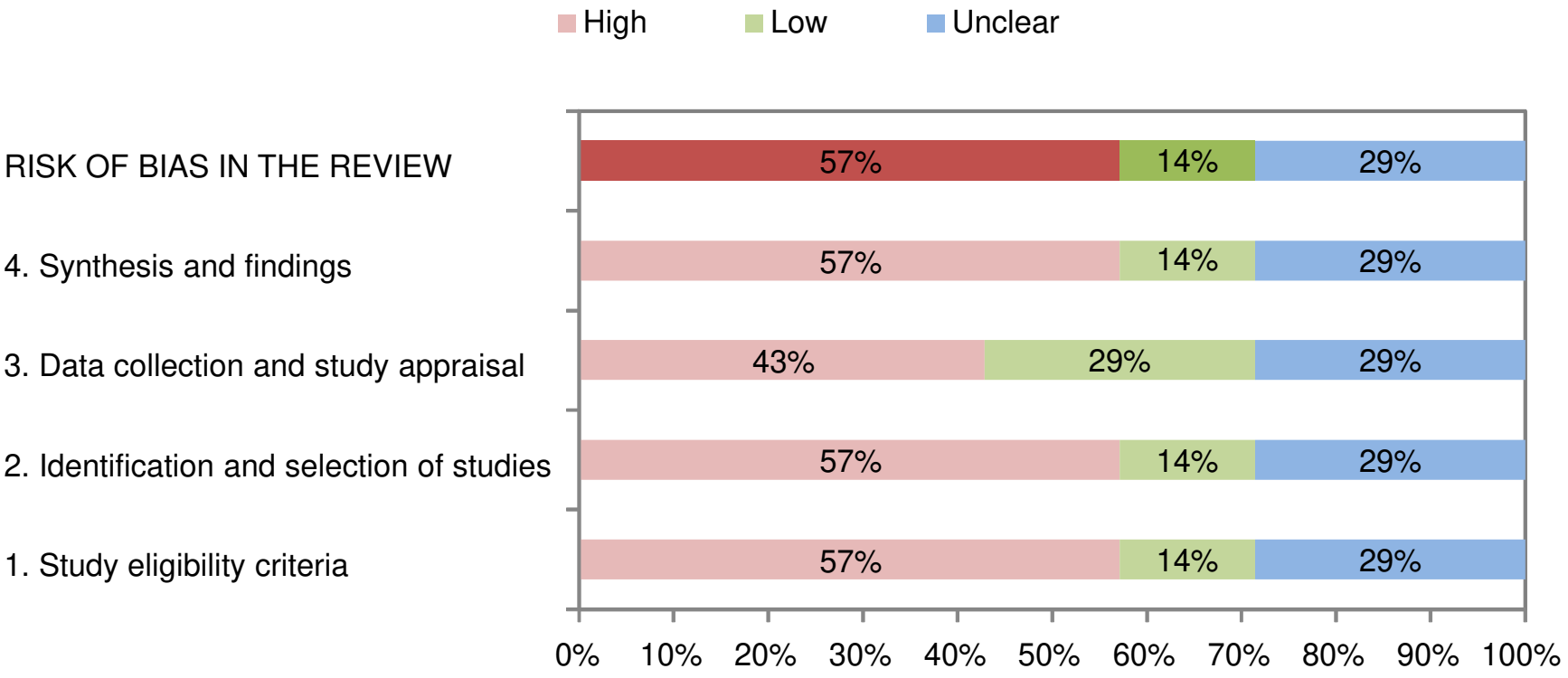


Fig. 3 Forest plot of all included meta-analyses of deep brain stimulation controlled trials on efficacy with standardized mean difference as type of metric. Summary of the result of all meta-analyses replicated from the included systematic reviews. Abbreviations: SMD, standardized mean difference; DBS, deep brain stimulation; YBOCS, Yale–Brown Obsessive Compulsive Scale.

Study	Comparison (left vs right)	Studies	Events /sample	I^2	SMD (95% CI)	95% predictive interval
Epilepsy						
Seizure freedom	Cerebellar DBS vs sham	3	-/33	0	-0.02 (-0.82 to 0.78)	-5.23 to 5.18
Seizure freedom	Hippocampal DBS vs sham	3	-/21	0	0.02 (-0.87 to 0.90)	-5.73 to 5.77
Responder rate	Cerebellar DBS vs sham	3	-/33	0	0.49 (-0.43 to 1.41)	-5.46 to 6.44
Responder rate	Hippocampal DBS vs sham	3	-/21	0	0.10 (-0.56 to 0.77)	-4.19 to 4.40
Obsessive compulsive disorder						
YBOCS	DBS vs sham	5	-/52	37	-0.95 (-1.74 to -0.15)	-3.12 to 1.23
Parkinson's disease						
> 50% reduction in seizure frequency	On vs off stimulation	1	115/499	-	0.10 (-0.36 to 0.57)	Not estimable

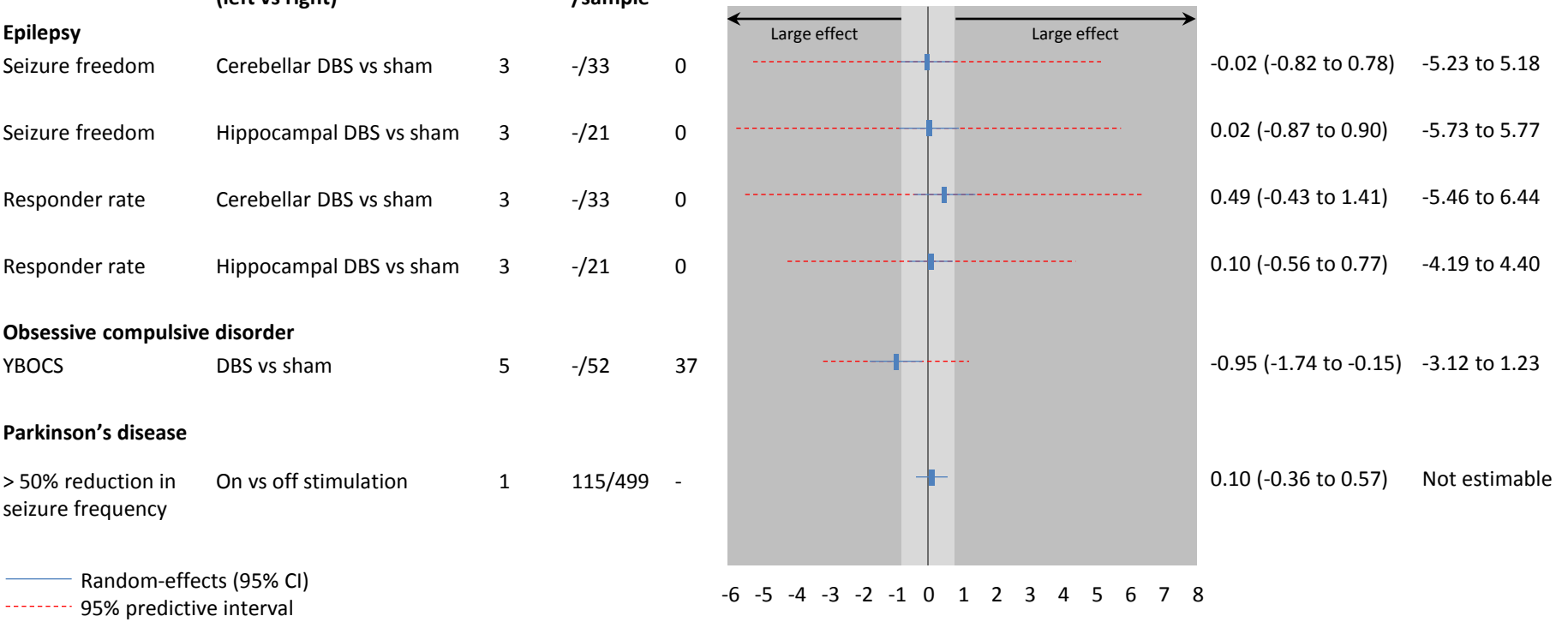
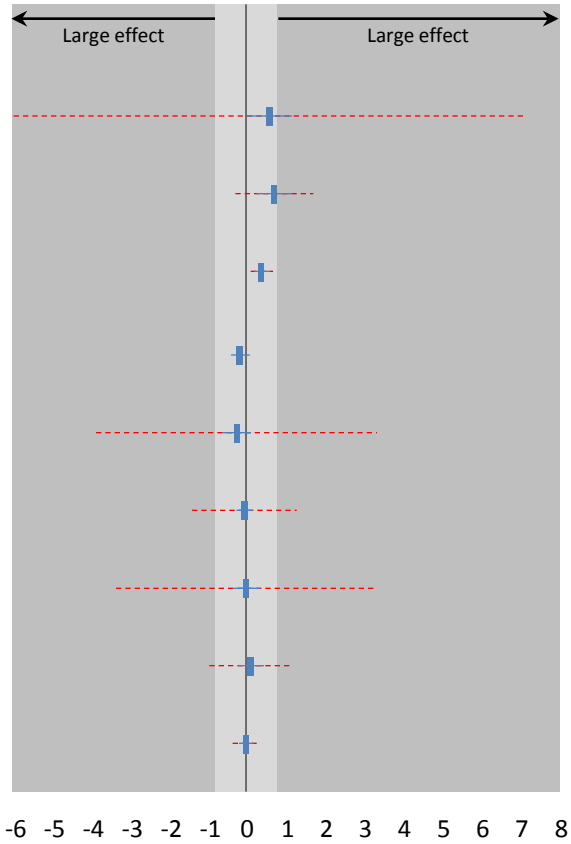


Fig. 4 Forest plot of all included meta-analyses of deep brain stimulation controlled trials on comparative effectiveness with standardized mean difference as type of metric. Summary of the result of all meta-analyses replicated from the included systematic reviews. Abbreviations: SMD, standardized mean difference; DBS, deep brain stimulation; STN, subthalamic nucleus; GPi, globus pallidus internus; BDI-II, Beck depression inventory-II; LED, levodopa equivalent dose; UPDRS, unified Parkinson's disease rating scale.

Study	Comparison (left vs right)	Studies	Events /sample	I^2	SMD (95% CI)	95% predictive interval
Parkinson's disease						
Depression	DBS (STN) vs DBS (GPi)	3	131/479	68	0.61 (0.03 to 1.19)	-5.94 to 7.16
Mortality	DBS (STN) vs DBS (GPi)	4	44/479	0	0.72 (0.27 to 1.18)	-0.27 to 1.72
BDI-II	DBS (STN) vs DBS (GPi)	5	-/540	0	0.39 (0.22 to 0.56)	0.12 to 0.67
LED	DBS (STN) vs DBS (GPi)	2	-/322	0	-0.16 (-0.38 to 0.06)	Not estimable
Phonemic verbal fluency	DBS (STN) vs DBS (GPi)	3	-/373	42	-0.24 (-0.60 to 0.12)	-3.84 to 3.35
Semantic verbal fluency	DBS (STN) vs DBS (GPi)	3	-/373	0	-0.05 (-0.25 to 0.16)	-1.38 to 1.29
UPDRS II (on-med)	DBS (STN) vs DBS (GPi)	3	-/450	52	0.01 (-0.32 to 0.33)	-3.33 to 3.35
UPDRS III (off-med)	DBS (STN) vs DBS (GPi)	5	-/518	60	0.12 (-0.22 to 0.45)	-0.94 to 1.17
UPDRS III (on-med)	DBS (STN) vs DBS (GPi)	5	-/518	17	0.01 (-0.18 to 0.19)	-0.34 to 0.35



Effectiveness and adverse effects of deep brain stimulation: umbrella review of meta-analyses

Tables

Table 1 Characteristics of the included systematic reviews

No	Review	Type	No of databases	Search period	Included trials	Outcomes	Mutliplicity addressed	Citations
1	Chambers 2013	Journal paper	4	Jan 2007-Dec 2012	RCT	1	-	10
2	Kisely 2014	Journal paper	2	-Apr 2013	RCT	1	-	7
3	Liu 2014	Journal paper	4	-Apr 2013	RCT	5	No	0
4	Sako 2014	Journal paper	1	1995-2013	RCT	7	No	8
5	Sprengers 2014	Cochrane Review	3	-Aug 2013	RCT	18	No	10
6	Arnaout 2015	Conference poster	1	-Nov 2014	non-RCT / RCT	5	No	-
7	Negida 2015	Conference poster	1	-Nov 2014	non-RCT / RCT	1	-	-

Abbreviations: RCT, randomized controlled trial.

Table 2 Tabular presentation for the ROBIS results of the included systematic reviews

Review	Phase 2				Phase 3
	1. study eligibility criteria	2. identification and selection of studies	3. data collection and study appraisal	4. synthesis and findings	risk of bias in the review
Chambers 2013	⊖	⊖	⊖	⊖	⊖
Kisely 2014	⊖	⊖	⊕	⊖	⊖
Liu 2014	⊖	⊖	⊖	⊖	⊖
Sako 2014	⊖	⊖	⊖	⊖	⊖
Sprengers 2014	⊕	⊕	⊕	⊕	⊕
Arnaout 2015	?	?	?	?	?
Negida 2015	?	?	?	?	?

Abbreviations: ⊕, low risk; ⊖, high risk; ?, unclear risk.

Table 3 Results of the included meta-analyses on efficacy (comparison of active interventions with sham controls)

No	First author / Year	Experimental	Control	Scope	Nature	Outcome	N	Cases/Sample	FE SMD (95%CI)	RE SMD (95%CI)	95% PrI	Largest SMD (95%CI)	I^2 (95% CI)¥	Egger P
	Epilepsy													
1	Sprengers 2014	Cerebellar DBS	Sham	Effect.	Bin	Seizure freedom	3	NR/33	-0.02 (-0.83,0.78)	-0.02 (-0.83,0.78)	-5.23,5.18	0.00 (-1.16,1.16)	0% (0%,73%)	-
2	Sprengers 2014	Hippocampal DBS	Sham	Effect.	Bin	Seizure freedom	3	NR/21	0.02 (-0.87,0.90)	0.02 (-0.87,0.90)	-5.73,5.77	0.00 (-1.30,1.30)	0% (0%,73%)	-
3	Sprengers 2014	Cerebellar DBS	Sham	Effect.	Bin	Responder rate	3	NR/33	0.49 (-0.43,1.41)	0.49 (-0.43,1.41)	-5.46,6.44	0.44 (-0.78,1.66)	0% (0%,73%)	-
4	Sprengers 2014	Hippocampal DBS	Sham	Effect.	Bin	Responder rate	3	NR/21	0.10 (-0.56,0.77)	0.10 (-0.56,0.77)	-4.19,4.40	0.00 (-0.81,0.81)	0% (0%,73%)	-
	OCD													
5	Kisely 2014	DBS	Sham	Effect.	Con	YBOCS	5	-/52	-0.87 (-1.47,-0.27)**	-0.95 (-1.74,-0.15)*	-3.12,1.23	-0.35 (-1.42,0.72)	37% (0%,76%)	-
	Parkinson's disease													
6	Chambers 2013	Stimulation on	Stimulation off	Effect.	Bin	> 50% Reduction in seizure frequency	1	115/499	0.10 (-0.36,0.57)	-	-	-	-	-

* p<0.05
** p<0.01
*** p<0.001
¥ significance results originating from the test for Q.
Abbreviations: FE, fixed-effect model; RE, random-effects model; N, number of studies; SMD, standardized mean difference; CI, confidence interval; PrI, predictive interval; DBS, deep brain stimulation; Bin, binary; Effect, effectiveness; NR, not reported; OCD, obsessive compulsive disorder; YBOCS, Yale–Brown Obsessive Compulsive Scale.

Table 4 Results of the included meta-analyses on comparative effectiveness (comparison of two active interventions)

No	First author / Year	Experimental	Control	Scope	Nature	Outcome	N	Cases/Sample	FE SMD (95%CI)	RE SMD (95%CI)	95% PrI	Largest SMD (95%CI)	I^2 (95% CI)¥	Egger P
	Parkinson's disease													
1	Sako 2014	DBS (STN)	DBS (GPi)	AE	Bin	Depression	3	131/479	0.41 (0.17,0.65)**	0.61 (0.03,1.19)*	-5.94,7.16	0.27 (0.00,0.54)	68% (0%,89%)*	-
2	Negida 2015	DBS (STN)	DBS (GPi)	AE	Bin	Mortality	4	44/479	0.72 (0.27, 1.18,)**	0.72 (0.27,1.18)**	-0.27,1.72	0.48 (-0.12,1.07)	0% (0%,68%)	-
3	Liu 2014	DBS (STN)	DBS (GPi)	Effect.	Con	BDI-II	5	-/540	0.39 (0.22,0.56)***	0.39 (0.22,0.56)***	0.12,0.67	0.30 (0.07,0.52)*	0% (0%,64%)	-
4	Liu 2014	DBS (STN)	DBS (GPi)	Effect.	Con	LED	2	-/322	-0.16 (-0.38,0.06)	-0.16 (-0.38,0.06)	NA	-0.18 (-0.40,0.05)	0% (NA)	-
5	Arnaout 2015	DBS (STN)	DBS (GPi)	Effect.	Con	Phonemic verbal fluency	3	-/373	-0.14 (-0.34,0.06)	-0.24 (-0.60,0.12)	-3.84,3.35	-0.06 (-0.28,0.16)	42% (0%,83%)	-
6	Arnaout 2015	DBS (STN)	DBS (GPi)	Effect.	Con	Semantic verbal fluency	3	-/373	-0.05 (-0.25,0.16)	-0.05 (-0.25,0.16)	-1.38,1.29	-0.01 (-0.24,0.22)	0% (0%,73%)	*
7	Liu 2014†	DBS (STN)	DBS (GPi)	Effect.	Con	UPDRS II (on-medication)	3	-/450	-0.04 (-0.23,0.14)	0.01 (-0.32,0.33)	-3.33,3.35	-0.18 (-0.40,0.05)	52% (0%,85%)	-
8	Liu 2014†	DBS (STN)	DBS (GPi)	Effect.	Con	UPDRS III (off-medication)	5	-/518	0.09 (-0.09,0.26)	0.12 (-0.23,0.46)	-0.94,1.17	-0.07 (-0.30,0.16)	60% (0%,83%)*	-
9	Liu 2014‡	DBS (STN)	DBS (GPi)	Effect.	Con	UPDRS III (on-medication)	5	-/518	0.00 (-0.17,0.17)	0.00 (-0.21,0.21)	-0.47,0.48	-0.07 (0.29,0.16)	17% (0%,70%)	-

* p<0.05
** p<0.01
*** p<0.001
†there was overlap with the poster report of Arnaout 2015, but the journal report was retained, due to completeness.
‡there was overlap with the poster report of Arnaout 2015, and the journal report of Sako 2014, but the journal report of Liu 2014, was retained, due to completeness and increased number of included studies.
¥ significance results originating from the test for Q.
Abbreviations: FE, fixed-effect model; RE, random-effects model; N, number of studies; SMD, standardized mean difference; CI, confidence interval; PrI, predictive interval; DBS, deep brain stimulation; Bin, binary; Effect, effectiveness; NA, not applicable; STN, subthalamic nucleus; GPi, globus pallidus internus; AE, adverse effects; BDI-II, Beck depression inventory-II; LED, levodopa equivalent dose; UPDRS, unified Parkinson's disease rating scale.

Effectiveness and adverse effects of deep brain stimulation: umbrella review of meta-analyses

Appendices

Appendix A Search strategies used for each database with the corresponding hits

Database	Search strategy	Limits	Hits	Duplicates	Unique
MEDLINE searched via PubMed (1950 –31.03.2015) www.ncbi.nlm.nih.gov/sites/entrez/	"deep brain stimulation"	Meta-analysis	28	0	28
Scopus (1966 - 31.03.2015) searched from www.scopus.com	"deep brain stimulation" AND ("systematic review" OR "meta-analysis")	Meta-analysis	49	14	35
CDSR & DARE searched via Cochrane Library on 31.03.2015 from www.thecochranelibrary.com	"deep brain stimulation"	Cochrane & other Reviews	23	14	9

Abbreviation: CDSR = Cochrane Database of Systematic Reviews; DARE = Database of Abstracts of Reviews of Effects.

Appendix B List of full text papers assessed for eligibility

Paper	Status
Brunoni AR, Amadera J, Berbel B, Volz MS, Rizzerio BG, Fregni F. A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. Int J Neuropsychopharmacol 2011;14:1133-45.	Excluded; not related to DBS
Ferreira JJ, Costa J, Coelho M, Sampaio C. The management of cervical dystonia. Expert Opin Pharmacother 2007;8:129-40.	Excluded; not related to DBS
Goetz CG, Poewe W, Rascol O, Sampaio C. Evidence-based medical review update: Pharmacological and surgical treatments of Parkinson's disease: 2001 to 2004. Mov Disord 2005;20:523-39.	Excluded; not related to DBS
Slotema CW, Blom JD, Van Lutterveld R, Hoek HW, Sommer IEC. Review of the efficacy of transcranial magnetic stimulation for auditory verbal hallucinations. Biol Psychiatry 2014;76:101-10.	Excluded; not related to DBS
Berlim MT, Neufeld NH, Eynde F. Repetitive transcranial magnetic stimulation (rTMS) for obsessive-compulsive disorder (OCD): an exploratory meta-analysis of randomized and sham-controlled trials. J Psychiatr Res 2013;8:999-1006.	Excluded; not related to DBS
Adeyemo BO, Simis M, Duarte Macea D, Fregni F. Systematic review of parameters of stimulation, clinical trial design characteristics, and motor outcomes in non-invasive brain stimulation in stroke. Front Psychiatry 2012;1:88.	Excluded; not related to DBS
Puig-Junoy J, Puig Peiro R. [Review of the economic evidence on the use of deep brain stimulation in late stage Parkinson's disease]. Neurologia (Barcelona, Spain) 2009;24:220-9.	Excluded; no quantitative synthesis
Dannon PN, Lowengrub K, Gonopolski Y, Kotler M. Current and emerging somatic treatment strategies in psychotic major depression. Expert Rev Med Devices 2006;6:73-80.	Excluded; no quantitative synthesis
Fasano A, Bove F, Lang AE. The treatment of dystonic tremor: A systematic review. J Neurol Neurosurg Psychiatry 2014;85:759-69.	Excluded; no quantitative synthesis
Grill WM. Safety considerations for deep brain stimulation: Review and analysis. Expert Rev Med Devices 2005;2:409-20.	Excluded; no quantitative synthesis
Gahr M, Connemann BJ, Freudenmann RW, Schonfeldt-Lecuona C. Safety of electroconvulsive therapy in the presence of cranial metallic objects. J ECT 2014;30:62-8.	Excluded; no quantitative synthesis
Takeshita S, Kurisu K, Trop L, Arita K, Akimitsu T, Verhoeff NP. Effect of subthalamic stimulation on mood state in Parkinson's disease: evaluation of previous facts and problems. Neurosurg Rev 2005;28:179-86.	Excluded; no quantitative synthesis
Temel Y, Kessels A, Tan S, Topdag A, Boon P, Visser-Vandewalle V. Behavioural changes after bilateral subthalamic stimulation in advanced Parkinson disease: a systematic review. Parkinsonism Relat Disord 2006;12:265-72.	Excluded; no quantitative synthesis
Vidailhet M, Jutras MF, Grabli D, Roze E. Deep brain stimulation for dystonia. J Neurol Neurosurg Psychiatry 2013;84:1029-42.	Excluded; no quantitative synthesis
Arumugham SS, Reddy JY. Augmentation strategies in obsessive-compulsive disorder. Expert Rev Neurother 2013;13:187-203.	Excluded; no quantitative synthesis
Fitzgerald PB. A review of developments in brain stimulation and the treatment of psychiatric disorders. Curr Psychiatry Rev 2006;2:199-205.	Excluded; no quantitative synthesis
Loo C, Katalinic N, Mitchell PB, Greenberg B. Physical treatments for bipolar disorder: A review of electroconvulsive therapy, stereotactic surgery and other brain stimulation techniques. J Affect Disord 2011;132:1-13.	Excluded; no quantitative synthesis

Hamani C, Pilitsis J, Rughani AI, Rosenow JM, Patil PG, Slavin KS, Abosch A, Eskandar E, Mitchell LS, Kalkanis S. Deep brain stimulation for obsessive-compulsive disorder: systematic review and evidence-based guideline sponsored by the American Society for Stereotactic and Functional Neurosurgery and the Congress of Neurological Surgeons (CNS) and endorsed by the CNS and American Association of Neurological Surgeons. Neurosurgery 2014;75:327-33.	Excluded; no quantitative synthesis
Department of Health, Australian Government. Deep brain stimulation for the symptoms of Parkinson's disease. Database of Abstracts of Reviews of Effects 2001;1:49.	Excluded; no quantitative synthesis
Romito LM, Albanese A. Dopaminergic therapy and subthalamic stimulation in Parkinson's disease: a review of 5-year reports. J Neurol 2010;Suppl2:S298-304.	Excluded; no quantitative synthesis
Nicholson T, Milne R. Pallidotomy, thalamotomy and deep brain stimulation for severe Parkinson's disease. Database of Abstracts of Reviews of Effects 1999;1:63.	Excluded; no quantitative synthesis
Jiang JL, Lo SF, Tsai ST, Chen SY. A systematic review of the impact of subthalamic nucleus stimulation on the quality of life of patients with Parkinson's disease. Database of Abstracts of Reviews of Effects 2014;1:15-20.	Excluded; no quantitative synthesis
Perestelo-Pérez L, Rivero-Santana A, Pérez-Ramos J, Serrano-Pérez P, Panetta J, Hilarion P. Deep brain stimulation in Parkinson's disease: meta-analysis of randomized controlled trials. J Neurol 2014´;261:2051-2060.	Excluded; no trial-specific data; requested from authors
Couto MI, Monteiro A, Oliveira A, Lunet N, Massano J. Depression and anxiety following Deep brain stimulation in Parkinson's disease: Systematic review and meta-analysis. Acta Med Port 2014;27:372-82.	Excluded; no trial-specific data; requested from authors
Appleby BS, Duggan PS, Regenberg A, Rabins PV. Psychiatric and neuropsychiatric adverse events associated with deep brain stimulation: A meta-analysis of ten years' experience. Mov Disord 2007;22:1722-8.	Excluded; not standard synthesis; averaging among studies
Bittar RG, Kar-Purkayastha I, Owen SL, Bear RE, Green A, Wang S, et al. Deep brain stimulation for pain relief: a meta-analysis. J Clin Neurosci 2005;12:515-9.	Excluded; not standard synthesis; averaging among studies
Stephen JH, Halpern CH, Barrios CJ, Balmuri U, Pisapia JM, Wolf JA, et al. Deep brain stimulation compared with methadone maintenance for the treatment of heroin dependence: a threshold and cost-effectiveness analysis. Addiction 2012;107:624-34.	Excluded; not standard synthesis; averaging among studies
Rughani AI, Lozano AM. Surgical treatment of myoclonus dystonia syndrome. Mov Disord 2013;3:282-7.	Excluded; not standard synthesis; averaging among studies
Andrade P, Carrillo-Ruiz JD, Jiménez F. A systematic review of the efficacy of globus pallidus stimulation in the treatment of Parkinson's disease. J Clin Neurosci 2009;16:877-81.	Excluded; not standard synthesis; averaging among studies followed by test between means
Woods SP, Rippeth JD, Conover E, Carey CL, Parsons TD, Troster AI. Statistical power of studies examining the cognitive effects of subthalamic nucleus deep brain stimulation in Parkinson's disease. Clin Neuropsychol 2006;20:27-38.	Excluded; not standard synthesis; averaging the power of included studies (no effect estimates)
Koy A, Hellmich M, Pauls KA, Marks W, Lin JP, Fricke O, et al. Effects of deep brain stimulation in dyskinetic cerebral palsy: a meta-analysis. Mov Disord 2013;28:647-54.	Excluded; not standard synthesis; correlation of single group estimates

Weaver F, Follett K, Hur K, Ippolito D, Stern M. Deep brain stimulation in Parkinson disease: a metaanalysis of patient outcomes. J Neurosurg 2005;103:956-67.	Excluded; not standard synthesis; meta-analysis of a single group estimate
Parsons TD, Rogers SA, Braaten AJ, Woods SP, Tröster AI. Cognitive sequelae of subthalamic nucleus deep brain stimulation in Parkinson's disease: a meta-analysis. Lancet Neurol 2006;5:578-88.	Excluded; not standard synthesis; meta-analysis of a single group estimate
Berlim MT, McGirr A, Van den Eynde F, Fleck MP, Giacobbe P. Effectiveness and acceptability of deep brain stimulation (DBS) of the subgenual cingulate cortex for treatment-resistant depression: a systematic review and exploratory meta-analysis. J Affect Disord 2014;159:31-8.	Excluded; not standard synthesis; meta-analysis of a single group estimate
Hoflich A, Savli M, Comasco E, Moser U, Novak K, Kasper S, et al. Neuropsychiatric deep brain stimulation for translational neuroimaging. Neuroimage 2013;79:30-41.	Excluded; not standard synthesis; meta-analysis of a single group estimate
Kimmelman J, Duckworth K, Ramsay T, Voss T, Ravina B, Emborg ME. Risk of surgical delivery to deep nuclei: a meta-analysis. Mov Disord 2011;26:1415-21.	Excluded; not standard synthesis; meta-analysis of a single group estimate
Boucai L, Cerquetti D, Merello M. Functional surgery for Parkinson's disease treatment: A structured analysis of a decade of published literature. Br J Neurosurg 2004;18:213-22.	Excluded; not standard synthesis; meta-analysis of a single group estimate
Kleiner-Fisman G, Herzog J, Fisman DN, Tamma F, Lyons KE, Pahwa R, et al. Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. Mov Disord 2006;21 Suppl 14:S290-304.	Excluded; not standard synthesis; meta-analysis of a single group estimate followed by meta-regression
Holloway KL, Baron MS, Brown R, Cifu DX, Carne W, Ramakrishnan V. Deep brain stimulation for dystonia: A meta-analysis. Neuromodulation 2006;9:253-61.	Excluded; not standard synthesis; meta-analysis of a single group estimate followed by meta-regression
St. George RJ, Nutt JG, Burchiel KJ, Horak FB. A meta-regression of the long-term effects of deep brain stimulation on balance and gait in PD. Neurology 2010;75:1292-9.	Excluded; not standard synthesis; meta-analysis of a single group estimate followed by meta-regression
Andrews C, Aviles-Olmos I, Hariz M, Foltynie T. Which patients with dystonia benefit from deep brain stimulation? A metaregression of individual patient outcomes.J Neurol Neurosurg Psychiatry 2010;81:1383-9.	Excluded; not standard synthesis; only meta-regression
Smith DF. Exploratory meta-analysis on deep brain stimulation in treatment-resistant depression. Acta Neuropsychiatr 2014;26:382-4.	Excluded; not standard synthesis; meta-analysis of a cohort studies with control data generated from field experts
Sako W, Miyazaki Y, Izumi Y, Kaji R. Which target is best for patients with Parkinson's disease? A meta-analysis of pallidal and subthalamic stimulation. J Neurol Neurosurg Psychiatry 2014;85:982-6.	Included; published paper
Chambers A, Bowen JM. Electrical stimulation for drug-resistant epilepsy: an evidence-based analysis. Ont Health Technol Assess Ser 2013;13:1-37.	Included; published paper
Liu Y, Li W, Tan C, Liu X, Wang X, Gui Y, et al. Meta-analysis comparing deep brain stimulation of the globus pallidus and subthalamic nucleus to treat advanced Parkinson disease. J Neurosurg 2014;121:709-18.	Included; published paper

Sprengers M, Vonck K, Carrette E, Marson AG, Boon P. Deep brain and cortical stimulation for epilepsy. Cochrane Database Syst Rev 2014;6:CD008497.	Included; published paper
Kisely S, Hall K, Siskind D, Frater J, Olson S, Crompton D. Deep brain stimulation for obsessive-compulsive disorder: A systematic review and meta-analysis. Psychol Med 2014;44:3533-42.	Included; published paper
Arnaout M, Negida A, El Ashal G, Fouda S, Ghanem E, El Ghonemy S. Meta-analysis comparing Subthalamic and Pallidal deep brain stimulation for patients with Parkinson's disease. Poster, ASIT conference 2015, Glasgow.	Included; conference poster
Negida A, Arnaout M, El Ashal G, Fouda S, Ghanem E, El Ghonemy S. Meta-analysis of mortality following Subthalamic and Pallidal deep brain stimulation for patients with Parkinson's disease. Poster, ASIT conference 2015, Glasgow.	Included; conference poster

Appendix C Assessments of risk of bias and methodological quality for the included systematic reviews

Assessment	Chambers 2013[32]	Kisely 2014[33]	Liu 2014[34]	Sako 2014[36]	Sprengers 2014[37]	Arnaout 2015[31]	Negida 2015[35]
Risk of bias of primary studies	Cochrane tool	Cochrane tool	Cochrane tool	-	Cochrane tool	-	-
Quality of evidence (GRADE approach)	Yes	Yes	-	-	Yes	-	-
Methodological adequacy of systematic review (AMSTAR tool)	7/11	5/11	7/11	2/11	11/11	NA	NA

Abbreviations: GRADE = Grading of Recommendations Assessment, Development and Evaluation; AMSTAR = A Measurement Tool to Assess Systematic Reviews; NA = not applicable.

Appendix D Assessment of the quality of the included systematic reviews with the AMSTAR tool

No	Item	Arnaout[1]	Chambers[2]	Kisely[3]	Liu[4]	Negida[5]	Sako[6]	Sprengers[7]
1	Was an 'a priori' design provided?	-	No mention	No mention	No mention	-	No mention	Yes (Cochrane Review)
2	Was there duplicate study selection and data extraction?	-	"Abstracts were reviewed by a single reviewer…"	"Data extraction was conducted by two independent researchers (J.H. and J.F.). All discrepancies during all stages of study selection, data extraction and quality assessment were resolved by re-checking source papers and further discussion among two other authors (S.K. and D.S.) to reach consensus."	"Two reviewers (Y.L. and C.T.) independently applied the inclusion and exclusion criteria, selected the studies, and extracted data and outcomes"	-	"Two authors doublechecked the inclusion criteria of the identified studies. … Two authors independently extracted data and checked each other. Any discrepancies were resolved by discussion."	"Four review authors (Mathieu Sprengers (MS),KristlVonck (KV), EvelienCarrette (EC) and Paul Boon (PB)) independently assessed the identified trials for inclusion."
3	Was a comprehensive literature search performed?	-	"..using Ovid MEDLINE, Ovid MEDLINE In-Process and Other Non-Indexed Citations, Ovid EMBASE, the Wiley Cochrane Library, and the Centre for Reviews and Dissemination database…"	"We conducted a comprehensive search using PubMed/Medline and EMBASE…"	"MEDLINE/PubMed, EMBASE, Web of Knowledge, and the Cochrane Library were searched…"	-	"We searched the PubMed database for publication…"	Three databases plus additional resources searched
4	Was the status of publication (i.e. grey literature) used as an inclusion criterion?	-	No mention	Posters, conference abstracts, and other unpublished reports were excluded	"Only published, English-language manuscripts were ultimately included in analyses"	-	No mention of grey literature, while the inclusion criteria include the word "publications"	"We contacted authors of relevant trials identified by our search, other researchers in the field, and manufacturers of the devices to identify unpublished or ongoing studies,…"
5	Was a list of studies (included and excluded) provided?	-	Not provided	Excluded studies not listed	Excluded studies not listed	-	Excluded studies not listed	"References to studies excluded from this review"
6	Were the characteristics of the included studies provided?	-	Table 3	Table 1	Table 1 (marginally extensive)	-	Characteristics missing from Table 1	"Characteristics of included studies"
7	Was the scientific quality of the included studies assessed and documented?	-	Table A4	"We assessed the quality of included studies using the following four criteria of the risk of bias assessment tool, developed by the Cochrane	"the same 2 reviewers in strict accordance with the Introduction to the Cochrane Handbook for Systematic Reviews of Interventions..."	-	No assessment	"The methodological quality of the studies was independently evaluated by two review authors (MS and KV) according to the guidelines in the Cochrane Handbook for Systematic Reviews of Interventions"

				Collaboration…"				
8	Was the scientific quality of the included studies used appropriately in formulating conclusions?	-	Conclusions accompanied with GRADE assessments	No mention	"Considering the limitations described above, caution should be taken in interpreting our findings"	-	No assessment	No specific mention of quality of evidence in the Conclusions or Implications sections
9	Were the methods used to combine the findings of studies appropriate?	-	Heterogeneity reported only in the forest plots; although no mention is made in the text, the authors include a comment about inconsistency in the Summary of Findings table, indicating that it was assessed	"We assessed heterogeneity using the I^2 statistic,…"	"Statistical analyses for continuous variables were performed, and heterogeneity was measured using I-square and chi-square tests."	-	"Random-effects models were employed for the meta-analysis because the underlying effect possibly differed across studies. The heterogeneity was assessed by a χ^2 test and designated as Q"	"Clinical heterogeneity was assessed by comparing the clinical and trial characteristics and a judgement was made as towhether significant clinical heterogeneity was present. Statistical inconsistency was assessed by visual inspection of the forest plots and by using the I² statistic and the Chi² test (Q test)."
10	Was the likelihood of publication bias assessed?	-	Assessed in table A1	"We were unable to test for publication bias as there were insufficient studies for any of the outcomes"	No mention	-	No mention	"As no more than three trials could be identified for each individual target, we were not able to assess the risk of publication bias."
11	Was the conflict of interest included?	-	Assessed in table A1	Conflicts of interest reported only for the review, not for the primary studies	Conflicts of interest reported only for the review, not for the primary studies	-	Conflicts of interest reported only for the review, not for the primary studies	Potential conflicts noted both for the primary studies and the review
	Sum	NA	7/11	5/11	7/11	NA	2/11	10/11

Abbreviation: NA = not applicable.

Appendix E Rating (Phase 2 and Phase 3) for each included systematic review according to the ROBIS tool

REVIEW: CHAMBERS 2013 (2)

Phase 2: Identifying concerns with the review process

DOMAIN 1: STUDY ELIGIBILITY CRITERIA					
Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
1.1 Did the review adhere to pre-defined objectives and eligibility criteria?				No mention of protocol or pre-defined criteria	
1.2 Were the eligibility criteria appropriate for the review question?				The primary research question entails information about the specific condition (drug-resistant epilepsy) and the patients (not surgical candidates) that were missing from the criteria	
1.3 Were eligibility criteria unambiguous?				The eligibility criteria were vague including the diagnosis modality, the treatments included and the technical details (stimulation magnitude, frequency, etc)	
1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?				Search started from 2007 on, as the relied on a previous HTA report that covered studies up to 2007. Certain studies however might have been missed.	
1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?		Probably ok, but no mention.			
	Low	**High**	**Unclear**		
Concerns regarding specification of study eligibility criteria		X			
Rationale for concern:		No existing protocol, non-specific eligibility criteria about the diagnostic method and the differentiation between surgical and non-surgical			

		candidates. Eligibility criteria were not clear, while the literature searched started only from 2007.			

DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES					
Describe methods of study identification and selection (e.g. number of reviewers involved):	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Authors searched MEDLINE, MEDLINE In-Process and Other Non-Indexed Citations, EMBASE, the Wiley Cochrane Library (although not specified which database), and the CRD.				
2.2 Were methods additional to database searching used to identify relevant reports?	Reference lists were checked.				
2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Full and appropriate search is provided in Appendix 1.				
2.4 Were restrictions based on date, publication format, or language appropriate?				English studies were excluded, which might introduce bias.	
2.5 Were efforts made to minimise error in selection of studies?				One reviewer screened abstracts and fulltexts.	
	Low	**High**	**Unclear**		
Concerns regarding methods used to identify and/or select studies		X			
Rationale for concern:		Restriction to studies in English and no efforts made to minimize error in study selection			

DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL					
Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
3.1 Were efforts made to minimise error in data collection?			No mention of data collection at all, but it is improbable that duplicate collection and appraisal was done.		
3.2 Were sufficient study characteristics available for both review authors and readers to be				Patient characteristics are inadequate	

able to interpret the results?				(diagnosis of epilepsy and degree). Study characteristics were also inadequate (single or multicenter).	
3.3 Were all relevant study results collected for use in the synthesis?			No mention on what data were collected and no mention on handling of missing data or data transformation.		
3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	The Cochrane tool was used.				
3.5 Were efforts made to minimise error in risk of bias assessment?			No mention of second reviewer performing or checking the risk of bias assessments.		
	Low	**High**	**Unclear**		
Concerns regarding methods used to collect data and appraise studies		X			
Rationale for concern:		Probably only a limited number of data were extracted from each trial, while it is improbable that error was controlled during risk of bias assessment			

DOMAIN 4: SYNTHESIS AND FINDINGS					
Describe synthesis methods:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
4.1 Did the synthesis include all studies that it should?			The study of Clarke et al was not used in the meta-analysis, due to missing outcome.		
4.2 Were all pre-defined analyses reported or departures explained?				No protocol.	
4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?				The author pooled studies with different interventions and with different experimental/control groups.	
4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Heterogeneity was assessed with I-square and Q test and no big evidence for inconsistency was found.				
4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?				No sensitivity or additional analyses conducted.	
4.6 Were biases in primary studies minimal or addressed in the synthesis?				Studies/outcomes with high risk of bias are included in the analyses, but no appropriate	

				measures have been taken to incorprorate this in the review's results accordingly.	
	Low	**High**	**Unclear**		
Concerns regarding the synthesis and findings LOW/HIGH/UNCLEAR		X			
Rationale for concern:		Missing data from Clarke et al., absence of protocol to compare departures from analysis plan, inappropriate meta-analyses (apples and oranges), no additional analyses to check for robustness and high risk of bias studies included without appropriate measures in the review.			

Phase 3: Judging risk of bias

Summarize the concerns identified during the Phase 2 assessment:

Domain	**Concern**	**Rationale for concern**
1. Concerns regarding specification of study eligibility criteria	High	No existing protocol, non-specific eligibility criteria about the diagnostic method and the differentiation between surgical and non-surgical candidates. Eligibility criteria were not clear, while the literature searched started only from 2007.
2. Concerns regarding methods used to identify and/or select studies	High	Restriction to studies in English and no efforts made to minimize error in study selection
3. Concerns regarding methods used to collect data and appraise studies	High	Probably only a limited number of data were extracted from each trial, while it is improbable that error was controlled during risk of bias assessment
4. Concerns regarding the synthesis and findings	High	Missing data from Clarke et al., absence of protocol to compare departures from

		analysis plan, inappropriate meta-analyses (apples and oranges), no additional analyses to check for robustness and high risk of bias studies included without appropriate measures in the review.

RISK OF BIAS IN THE REVIEW					
Describe whether conclusions were supported by the evidence:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
A. Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?				Not all issues are addressed in the review.	
B. Was the relevance of identified studies to the review's research question appropriately considered?				Only part of the research questions were answered from the included studies. Although, the effectiveness of DBS and VNS was the scope of the review, many studies included no untreated control and only compared low to high stimulation, which is no direct effectiveness measure.	
C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?				No, the authors focused mainly on the statistical significance of the findings.	
	Low	**High**	**Unclear**		
Risk of bias in the review risk:		X			
Rationale for risk:		Many residual issues regarding all methodological aspects of the review that were not taken into account in the review's conclusions. Additionally, most studies compared only a relative effectiveness between low and high stimulation, but no absolute effectiveness compared to no treatment.			

Phase 2: Identifying concerns with the review process

DOMAIN 1: STUDY ELIGIBILITY CRITERIA					
Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
1.1 Did the review adhere to pre-defined objectives and eligibility criteria?				No relevant information provided.	
1.2 Were the eligibility criteria appropriate for the review question?		Eligibility criteria probably appropriate. The patient component is vague and not good defined, but this probably lies on the broad question of the review.			
1.3 Were eligibility criteria unambiguous?	Eligibility criteria clearly-defined. Some ambiguitiy regarding the inclusion of various study designs, but the authors acknowledge it and take appropriate steps.				
1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	No restrictions imposed.				
1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	No restrictions imposed.				
	Low	**High**	**Unclear**		
Concerns regarding specification of study eligibility criteria		X			
Rationale for concern:		No evidence of a priori design for the review procedures and this is difficult to judge, due to missing pertinent information.			

X

DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES					
Describe methods of study identification and selection (e.g. number of reviewers involved):	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished			Only PubMed and Embase were searched. These two databases have		

reports?			large overlap, limited coverage and very little gray or non-English literature.		
2.2 Were methods additional to database searching used to identify relevant reports?				No mention.	
2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?					The full search strategy was not reported and there were no details of the search terms; there was therefore no information on which to base the assessment for this question.
2.4 Were restrictions based on date, publication format, or language appropriate?	No restrictions mentioned.				
2.5 Were efforts made to minimise error in selection of studies?				No mention of duplicate procedures—study selection was probably done from a single reviewer without error management.	
	Low	**High**	**Unclear**		
Concerns regarding methods used to identify and/or select studies		X			
Rationale for concern:		Limited literature search, without additional methods to maximize identification of studies, incomplete reporting and high risk of error during the study selection.			

DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL					
Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
3.1 Were efforts made to minimise error in data collection?	Procedures performed independently by two authors.				
3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Data collected and reported in Table 1 seem adequate (study design, patients allocated-finished, DBS information, and outcomes reported).				
3.3 Were all relevant study results collected for use in the synthesis?		Some information about the patient samples are missing			

		from some studies, probably due to incomplete reporting from the studies.			
3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	The Cochrane tool was used.				
3.5 Were efforts made to minimise error in risk of bias assessment?	Procedures performed independently by two authors.				
	Low	**High**	**Unclear**		
Concerns regarding methods used to collect data and appraise studies	**X**				
Rationale for concern:	Data collection and risk of bias assessment done under management of probable errors. Data collected are probably adequate, while an appropriate risk-of-bias tool was used (Cochrane tool).				

DOMAIN 4: SYNTHESIS AND FINDINGS					
Describe synthesis methods:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
4.1 Did the synthesis include all studies that it should?	All studies are included (flowdiagrama vs forest plot).				
4.2 Were all pre-defined analyses reported or departures explained?				No protocol existed.	
4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Appropriate comparisons with similar PICOS. Both changes in end measurements and Pre-Post increments are pooled. This has been shown to be appropriate through empirical studies. The authors, assessed also these two measures independently.				
4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Tau-squared moderate. I-square and Q don't indicate excessive inconsistency.				
4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Sensitivity and subgroup analyses were planned and conducted, but were not always possible.				
4.6 Were biases in primary studies minimal or addressed in the synthesis?			The authors acknowledge that outcomes were missing from 80% of		

			the included studies.		
	Low	**High**	**Unclear**		
Concerns regarding the synthesis and findings LOW/HIGH/UNCLEAR		X			
Rationale for concern:		Missing outcomes from 4/5 of the included studies and no existing protocol to judge any possible departures from analysis plan.			

Phase 3: Judging risk of bias

Summarize the concerns identified during the Phase 2 assessment:

Domain	**Concern**	**Rationale for concern**
1. Concerns regarding specification of study eligibility criteria	High	No evidence of a priori design for the review procedures and this is difficult to judge, due to missing pertinent information.
2. Concerns regarding methods used to identify and/or select studies	High	Limited literature search, without additional methods to maximize identification of studies, incomplete reporting and high risk of error during the study selection.
3. Concerns regarding methods used to collect data and appraise studies	Low	Data collection and risk of bias assessment done under management of probable errors. Data collected are probably adequate, while an appropriate risk-of-bias tool was used (Cochrane tool).
4. Concerns regarding the synthesis and findings	High	Missing outcomes from 4/5 of the included studies and no existing protocol to judge any possible departures from analysis plan.

RISK OF BIAS IN THE REVIEW					
Describe whether conclusions were supported by the evidence:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
A. Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?				Not all concerns addressed.	
B. Was the relevance of identified studies to the review's research	Relevant studies for the review's				

question appropriately considered?	question				
C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?	Authors clearly acknowledge that significance testing might have been confounded by limited samples.				
	Low	**High**	**Unclear**		
Risk of bias in the review risk:		**X**			
Rationale for risk:		Absence of protocol to judge a priori design and no mention of pre-defined criteria. Limited literature search and risk of error during study selection. Missing outcomes in 4/5 included studies and authors' judgement about possible risk of publication bias.			

SR: LIU 2014 (4)

Phase 2: Identifying concerns with the review process

DOMAIN 1: STUDY ELIGIBILITY CRITERIA					
Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
1.1 Did the review adhere to pre-defined objectives and eligibility criteria?				No protocol provided.	
1.2 Were the eligibility criteria appropriate for the review question?	Yes, the eligibility criteria were relatively clear and, in any way, more specific than the research question. All PICOS components are covered.				
1.3 Were eligibility criteria unambiguous?	Eligibility criteria clear enough.				
1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	No pertinent eligibility criteria.				
1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?				Only published and English language studies were included.	
	Low	**High**	**Unclear**		

Concerns regarding specification of study eligibility criteria		**X**			
Rationale for concern:		The systematic review was not registered, while only English and published studies were included, which can introduce bias.			

DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES					
Describe methods of study identification and selection (e.g. number of reviewers involved):	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes; MEDLINE, Embase, Web of Knowledge and Cochrane Library (not specified which sub-library) were searched.				
2.2 Were methods additional to database searching used to identify relevant reports?	Reference lists were checked.				
2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?		The complete search strategy is not provided; just keywords that were used and no obvious reason to assume inappropriateness exists.			
2.4 Were restrictions based on date, publication format, or language appropriate?				Only published and English language studies were included.	
2.5 Were efforts made to minimise error in selection of studies?	Duplicate study selection was performed.				
	Low	**High**	**Unclear**		
Concerns regarding methods used to identify and/or select studies		**X**			
Rationale for concern:		Only published and English language studies were included.			

DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL					
Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
3.1 Were efforts made to minimise error in data collection?	Duplicate data collection was performed.				
3.2 Were sufficient study characteristics available for both			Some data missing from the included		

review authors and readers to be able to interpret the results?			studies.		
3.3 Were all relevant study results collected for use in the synthesis?		Some data were missing: study design (single- or multicenter), study setting, severity of disease, technical details of the intervention.			
3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes (Cochrane tool was used.				
3.5 Were efforts made to minimise error in risk of bias assessment?	Duplicate risk of bias assessment was performed.				
	Low	**High**	**Unclear**		
Concerns regarding methods used to collect data and appraise studies		**X**			
Rationale for concern:		Some data missing from the included studies.			

DOMAIN 4: SYNTHESIS AND FINDINGS					
Describe synthesis methods:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
4.1 Did the synthesis include all studies that it should?				Incomplete reporting from the included studies meant that many studies were omitted from the analyses. Half of the trials on adverse effects reported incomplete data.	
4.2 Were all pre-defined analyses reported or departures explained?				No protocol available. The post hoc model choice indicates that no a priori analysis plan was made.	
4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Analyses were appropriate.				
4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Moderate heterogeneity was identified and was assessed with sensitivity analyses.				
4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	All sensitivity analyses that were performed (by omission of a single study) indicated robustness.				
4.6 Were biases in primary studies minimal or addressed in the synthesis?			A trial had high risk of bias for one domain, but this was not incorporated in synthesis. Otherwise the included trials		

			had unclear risk of bias for many domains.		
	Low	**High**	**Unclear**		
Concerns regarding the synthesis and findings LOW/HIGH/UNCLEAR		**X**			
Rationale for concern:		Amon the 11 excluded studies were also studies with non-extractable results. Half of the trials on adverse effects reported incomplete data. No protocol available. The post hoc model choice indicates that no a priori analysis plan was made.			

Phase 3: Judging risk of bias

Summarize the concerns identified during the Phase 2 assessment:

Domain	**Concern**	**Rationale for concern**
1. Concerns regarding specification of study eligibility criteria	High	The systematic review was not registered, while only English and published studies were included, which can introduce bias.
2. Concerns regarding methods used to identify and/or select studies	High	Only published and English language studies were included.
3. Concerns regarding methods used to collect data and appraise studies	High	Some data missing from the included studies.
4. Concerns regarding the synthesis and findings	High	Amon the 11 excluded studies were also studies with non-extractable results. Half of the trials on adverse effects reported incomplete data. No protocol available. The post hoc model choice indicates that no a priori analysis plan was made.

RISK OF BIAS IN THE REVIEW					
Describe whether conclusions were supported by the evidence:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
A. Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?			English language and missing data were addressed by the authors of the		

			review. Absence of gray literature and missing a priori design for the review and analysis plan are not reported.		
B. Was the relevance of identified studies to the review's research question appropriately considered?	Judged as appropriate.				
C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?				Interpretation is based mainly on the statistical signifcance of the overall pooled estimate. For many forest plots, the included trials provide contradictory results.	
	Low	**High**	**Unclear**		
Risk of bias in the review risk:		**X**			
Rationale for risk:		Absence of gray literature and missing a priori design for the review and analysis plan are not reported. Interpretation is based mainly on the statistical signifcance of the overall pooled estimate. For many forest plots, the included trials provide contradictory results.			

ROBIS TOOL – SR: SAKO 2014 (6)

Phase 2: Identifying concerns with the review process

DOMAIN 1: STUDY ELIGIBILITY CRITERIA					
Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
1.1 Did the review adhere to pre-defined objectives and eligibility criteria?				No protocol provided.	
1.2 Were the eligibility criteria appropriate for the review question?		No clear research question. The eligibility criteria however were clear enough and would match the general framework of the question.			

1.3 Were eligibility criteria unambiguous?	The eligibility criteria were clear.				
1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?				Search starting date was set at 1995 without justification. Setting a minimum of 10 patients seems logical.	
1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?				It is not explicitely stated that only published papers were included, but MEDLINE that was searched, indexes manily published reports. The authors also included only English papers.	
	Low	**High**	**Unclear**		
Concerns regarding specification of study eligibility criteria		X			
Rationale for concern:		No protocol provided, search limited to studies published after 1995 and only in English language.			

DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES					
Describe methods of study identification and selection (e.g. number of reviewers involved):	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?				Only MEDLINE was searched.	
2.2 Were methods additional to database searching used to identify relevant reports?				No additional search methods. Four additional papers were found from existing review articles, but this was not judged as an appropriate adjunct to database search.	
2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?		No complete search strategy reported, but the provided terms seem appropriate.			
2.4 Were restrictions based on date, publication format, or language appropriate?				Inappropriate restrictions regarding date, language, and (possibly) format.	
2.5 Were efforts made to minimise error in selection of studies?	Two reviewers selected / checked eligible studies.				
	Low	**High**	**Unclear**		
Concerns regarding methods used to identify and/or select studies		X			
Rationale for concern:		Limited database			

		search and no additional search methods, together with inappropriate restrictions, might have resulted in missing a potentially eligible trial.			

DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL					
Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
3.1 Were efforts made to minimise error in data collection?	Duplicate procedure by two authors.				
3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?				Missing information about the design and setting of included trials. No data about diagnosis and severity of the disease extracted.	
3.3 Were all relevant study results collected for use in the synthesis?				Missing data from trials in Table 1.	
3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?				Not risk of bias assessment.	
3.5 Were efforts made to minimise error in risk of bias assessment?				Not risk of bias assessment.	
	Low	**High**	**Unclear**		
Concerns regarding methods used to collect data and appraise studies		X			
Rationale for concern:		Limited data collected from the included trials, incomplete data reporting from the included trials and formal assessment of risk of bias.			

DOMAIN 4: SYNTHESIS AND FINDINGS					
Describe synthesis methods:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
4.1 Did the synthesis include all studies that it should?	All trials included in the meta-analysis.				
4.2 Were all pre-defined analyses reported or departures explained?				No protocol provided.	
4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Judged as appropriate.				
4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?		Moderate inconsistency detected, while			

		direction of effects is unanimous.			
4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?				No additional analyses conducted.	
4.6 Were biases in primary studies minimal or addressed in the synthesis?				Biases were not addressed in general.	
	Low	**High**	**Unclear**		
Concerns regarding the synthesis and findings LOW/HIGH/UNCLEAR		X			
Rationale for concern:		Absence of a pre-defined analysis plan, while additional analyses and sensitivity analyses were not conducted.			

Phase 3: Judging risk of bias

Summarize the concerns identified during the Phase 2 assessment:

Domain	**Concern**	**Rationale for concern**
1. Concerns regarding specification of study eligibility criteria	High	No protocol provided, search limited to studies published after 1995 and only in English language.
2. Concerns regarding methods used to identify and/or select studies	High	Limited database search and no additional search methods, together with inappropriate restrictions, might have resulted in missing a potentially eligible trial.
3. Concerns regarding methods used to collect data and appraise studies	High	Limited data collected from the included trials, incomplete data reporting from the included trials and formal assessment of risk of bias.
4. Concerns regarding the synthesis and findings	High	Absence of a pre-defined analysis plan, while additional analyses and sensitivity analyses were not conducted.

RISK OF BIAS IN THE REVIEW					
Describe whether conclusions were supported by the evidence:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
A. Did the interpretation of findings address all of the concerns				No; most of the concerns remained	

identified in Domains 1 to 4?				unadressed by the authors.	
B. Was the relevance of identified studies to the review's research question appropriately considered?	Probably yes, as the research question was very broad.				
C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?				No; statistical significance was the main factor of focus.	
	Low	**High**	**Unclear**		
Risk of bias in the review risk:		**X**			
Rationale for risk:		Residual concerns that were identified in Phase 2 and remained unaddressed.			

REVIEW: SPRENGERS 2014 (7)

Phase 2: Identifying concerns with the review process

DOMAIN 1: STUDY ELIGIBILITY CRITERIA					
Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
1.1 Did the review adhere to pre-defined objectives and eligibility criteria?	Yes, as it was a Cochrane Review.				
1.2 Were the eligibility criteria appropriate for the review question?	Yes, criteria properly fit the question.				
1.3 Were eligibility criteria unambiguous?	Eligibility criteria were clear and specific.				
1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	No restrictions imposed.				
1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	No restrictions imposed.				
	Low	**High**	**Unclear**		
Concerns regarding specification of study eligibility criteria	**X**				
Rationale for concern:	No concern				

DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES					
Describe methods of study identification and selection (e.g. number of reviewers involved):	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
2.1 Did the search include an appropriate range of databases/electronic sources for	MEDLINE, Cochrane special register from the				

published and unpublished reports?	Epilepsy group and CENTRAL.				
2.2 Were methods additional to database searching used to identify relevant reports?	Reference lists, relevant and other authors' contact, and manufacturers were investigated.				
2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Judged as appropriate.				
2.4 Were restrictions based on date, publication format, or language appropriate?	None imposed.				
2.5 Were efforts made to minimise error in selection of studies?	Selection by four authors independently.				
	Low	**High**	**Unclear**		
Concerns regarding methods used to identify and/or select studies	X				
Rationale for concern:	No concern.				

DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL					
Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
3.1 Were efforts made to minimise error in data collection?	Extraction from two authors independently.				
3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Adequate data extracted.				
3.3 Were all relevant study results collected for use in the synthesis?	Adequate data were collected. In case of missing data, authors were contacted and/or the missing data were calculated.				
3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	With the Cochrane risk of bias tool.				
3.5 Were efforts made to minimise error in risk of bias assessment?	Procedure conducted in duplicate.				
	Low	**High**	**Unclear**		
Concerns regarding methods used to collect data and appraise studies	X				
Rationale for concern:	No concern.				

DOMAIN 4: SYNTHESIS AND FINDINGS					

Describe synthesis methods:	Yes	Probably Yes	Probably No	No	No Information
4.1 Did the synthesis include all studies that it should?	Apparently yes.				
4.2 Were all pre-defined analyses reported or departures explained?	All departures explained and justified.				
4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?		Synthesis appropriate in general terms, although model choice was based on post hoc availability of studies and post hoc inspection of statistical heterogeneity, which is not a robust justification. Clinical and statistical reasoning should be, at least in part, provided a priori.			
4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Both clinical and statistical heterogeneity were assessed (the latter with I-square and Q).				
4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Subgroup analyses were planned and performed. Sensitivity analyses were planned, but could not be successfully conducted.				
4.6 Were biases in primary studies minimal or addressed in the synthesis?	Fully incorporated both with the GRADE approacha and by consistent mentioning in the limitations and conclusions of the review.				
	Low	**High**	**Unclear**		
Concerns regarding the synthesis and findings LOW/HIGH/UNCLEAR	**X**				
Rationale for concern:	No (or only minor) concerns.				

Phase 3: Judging risk of bias

Summarize the concerns identified during the Phase 2 assessment:

Domain	Concern	Rationale for concern
1. Concerns regarding specification of study eligibility criteria	Low	No concerns
2. Concerns regarding methods used to identify and/or select studies	Low	No concerns

3. Concerns regarding methods used to collect data and appraise studies	Low	No concerns
4. Concerns regarding the synthesis and findings	Low	No (or only minor) concerns

RISK OF BIAS IN THE REVIEW					
Describe whether conclusions were supported by the evidence:	**Yes**	**Probably Yes**	**Probably No**	**No**	**No Information**
A. Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?	Fully incorporated.				
B. Was the relevance of identified studies to the review's research question appropriately considered?	Direct relevance.				
C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?	Interpretation together with associated uncertainty and possible confounders.				
	Low	**High**	**Unclear**		
Risk of bias in the review risk:	X				
Rationale for risk:	No concerns				

Appendix F Included studies' matrix from meta-analyses with overlap

Nature	Effectiveness																										Adverse effects													
Outcome	UPDRS II						UPDRS III_off							UPDRS III_on							LED		semantic			phonetic			≥50% seizure reduction	BDI-II					Depression					Mortality			
Trial	1	2	7	8	10	11	2	5	7	8	9	10	11	2	5	7	8	9	10	11	2	7	3	4	7	3	4	7	6	3	5	7	9	10	2	5	7	9	10	2	4	8	11
Review																																											
Chambers 2013 (2)																													x														
Liu 2014 (4)		x	x		x		x	x	x		x	x		x	x	x		x	x		x	x								x	x	x	x	x								
Sako 2014 (6)							x	x	x			x		x	x	x			x																	x	x		x				
Arnaout 2015 (1)	x	x		x	x	x	x	x		x	x	x	x	x	x		x	x	x	x			x	x	x	x	x	x														
Negida 2015 (5)																																									x	x	x	x

Trial number	Reference
1	Krause M, Fogel W, Heck A, et al. Deep brain stimulation for the treatment of Parkinson's disease: subthalamic nucleus versus globus pallidus internus. *J Neurol Neurosurg Psychiatry* 2001;70:464-70.
2	Anderson VC, Burchiel KJ, Hogarth P, Favre J, Hammerstad JP. Pallidal vs subthalamic nucleus deep brain stimulation in Parkinson disease. *Arch Neurol* 2005;62:554-60.
3	Rothlind JC, Cockshott RW, Starr PA, Marks WJ. Neuropsychological performance following staged bilateral pallidal or subthalamic nucleus deep brain stimulation for Parkinson's disease. *J Int Neuropsychol Soc* 2007;13:68-79.
4	Okun MS, Fernandez HH, Wu SS, et al. Cognition and mood in Parkinson's disease in subthalamic nucleus versus globus pallidus interna deep brain stimulation: the COMPARE trial. *Ann Neurol* 2009;65:586-95.
5	Zahodne LB, Okun MS, Foote KD, et al. Greater improvement in quality of life following unilateral deep brain stimulation surgery in the globus pallidus as compared to the subthalamic nucleus. *J Neurol* 2009;256:1321-9.
6	Fisher R, Salanova V, Witt T, et al. Electrical stimulation of the anterior nucleus of thalamus for treatment of refractory epilepsy. *Epilepsia* 2010;51:899-908.
7	Follett KA, Weaver FM, Stern M, et al. Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. *N Engl J Med* 2010;362:2077-91.
8	Moro E, Lozano AM, Pollak P, et al. Long-term results of a multicenter study on subthalamic and pallidal stimulation in Parkinson's disease. *Mov Disord* 2010;25:578-86.
9	Rocchi L, Carlson-Kuhta P, Chiari L, Burchiel KJ, Hogarth P, Horak FB. Effects of deep brain stimulation in the subthalamic nucleus or globus pallidus internus on step initiation in Parkinson disease: laboratory investigation. *J Neurosurg* 2012;117:1141-9.
10	Odekerken VJJ, van Laar T, Staal MJ, et al. Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. *Lancet Neurol* 2013;12:37-44.
11	Weaver FM, Follett KA, Stern M, et al. Randomized trial of deep brain stimulation for Parkinson disease: thirty-six-month outcomes. *Neurology* 2012;79:55-65.

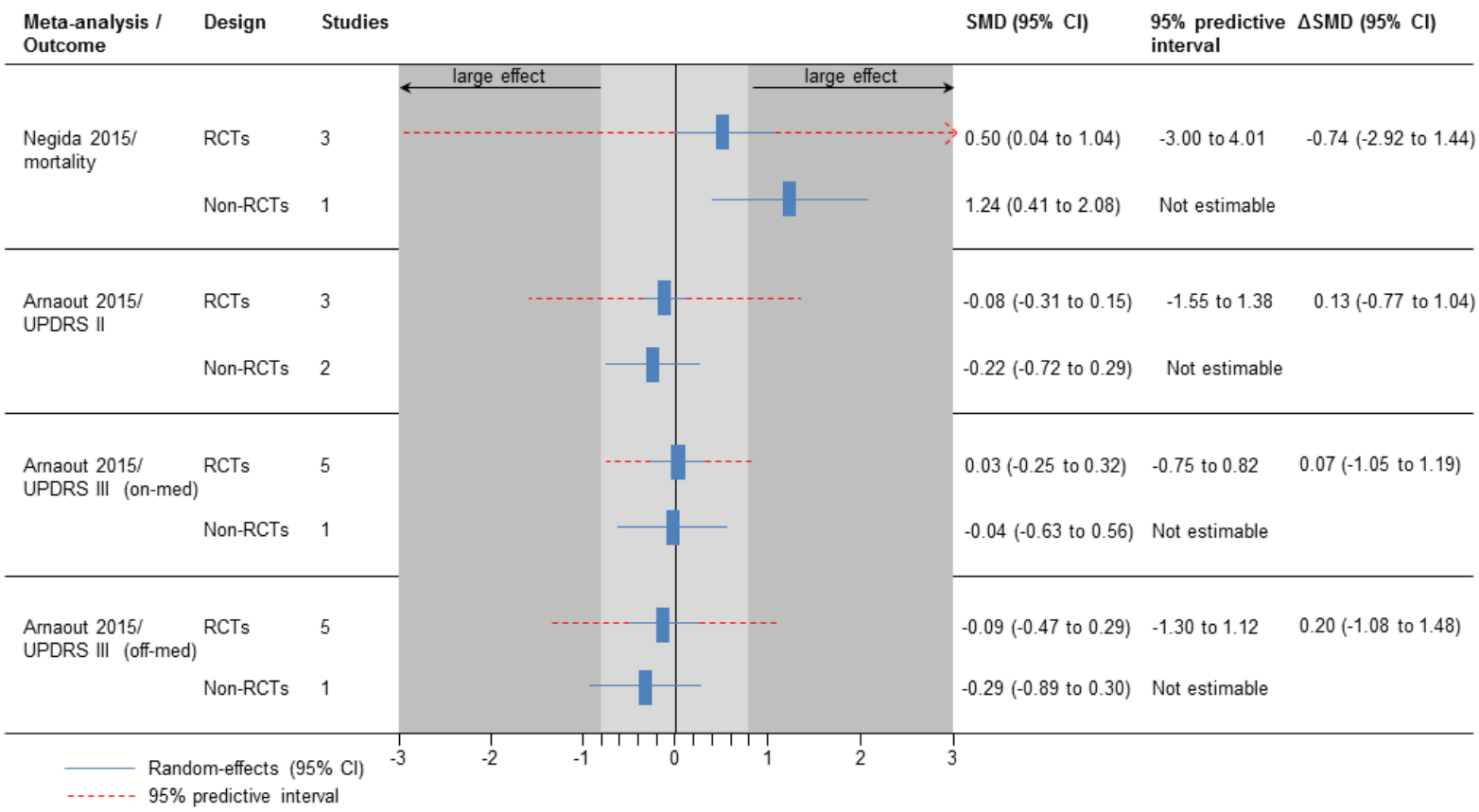
Abbreviations: UPDRS = unified Parkinson's disease rating scale; LED = levodopa equivalent dose; BDI-II = Beck depression inventory-II.

Appendix G Results from meta-analyses with overlap

Outcome	Meta-analysis	N	Cases/sample	FE	RE	95% PrI	Largest trial	I^2 (95% CI)	Egger
UPDRS II	Liu[4] (included)	3	NA/450	-0·04 (-0·23,0·14)	0·01 (-0·32,0·33)	-3·33,3·35	-0·18 (-0·40,0·05)	52% (0%,85%)	-
	Arnaout[1] (omitted)	5	NA/373	-0·10 (-0·31,0·10)	-0·10 (-0·31,0·10)	-0·44,0·23	0·02 (-0·29,0·33)	0% (0%,64%)	-
UPDRS III (on-medication)	Liu[4] (included)	5	NA/518	0·09 (-0·09,0·26)	0·12 (-0·23,0·46)	-0·94,1·17	-0·07 (-0·30,0·16)	60% (0%,83%)*	-
	Arnaout[1] (excluded)	6	NA/426	0·04 (-0·15,0·23)	0·03 (-0·21,0·26)	-0·49,0·55	0·32 (0·01,0·64)	24% (0%,70%)	-
UPDRS III (off-medication)	Liu[4] (included)	5	NA/518	0·00 (-0·17,0·17)	0·00 (-0·21,0·21)	-0·47,0·48	-0·07 (0·29,0·16)	17% (0%,70%)	-
	Arnaout[1] (omitted)	6	NA/426	-0·11 (-0·30,0·09)	-0·12 (-0·44,0·20)	-1·04,0·79	0·24 (-0·07,0·55)	56% (0%,80%)*	*
	Sako[6] (omitted)	4	-	0·12 (-0·07,0·30)	0·19 (-0·20,0·58)	-1·42,1·80	-0·09 (-0·33,0·15)	68% (0%,87%)	-

Abbreviations: FE = fixed-effect model; RE = random-effects model; PrI = predictive interval; CI = confidence interval; UPDRS = unified Parkinson's disease rating scale; NA = not applicable.
*significant at the 10% level.

Appendix H Results of the sensitivity analysis according to basic study design



.

Appendix I Results from meta-analyses with overlap

Meta-analysis	P<0·001	Adequate sample	I^2<75%	95% predictive interval excludes null	No Egger P<0·10	Meeting all criteria
1	No	No	Yes	No	Yes	No
2	No	No	Yes	No	Yes	No
3	No	No	Yes	No	Yes	No
4	No	No	Yes	No	Yes	No
5	No	No	Yes	No	Yes	No
6	No	No	Yes	No	Yes	No
7	No	No	Yes	No	Yes	No
8	Yes	No	Yes	Yes	Yes	No
9	No	No	Yes	No	Yes	No
10	No	No	Yes	No	Yes	No
11	No	No	Yes	No	Yes	No
12	No	No	Yes	No	No	No
13	No	No	Yes	No	Yes	No
14	No	No	Yes	No	Yes	No
15	No	No	Yes	No	Yes	No

References

1. Arnaout M, Negida A, El Ashal G, Fouda S, Ghanem E, El Ghonemy S. Meta-analysis comparing Subthalamic and Pallidal deep brain stimulation for patients with Parkinson's disease. Poster, Association of Surgeons in Training conference, 27 Feb-1 March 2015, Glasgow.
2. Chambers A, Bowen JM. Electrical stimulation for drug-resistant epilepsy: an evidence-based analysis. Ont Health Technol Assess Ser 2013;13:1–37.
3. Kisely S, Hall K, Siskind D, Frater J, Olson S, Crompton D. Deep brain stimulation for obsessive-compulsive disorder: A systematic review and meta-analysis. Psychol Med 2014;44:3533–42.
4. Liu Y, Li W, Tan C, et al. Meta-analysis comparing deep brain stimulation of the globus pallidus and subthalamic nucleus to treat advanced Parkinson disease. J Neurosurg 2014;121:709–18.
5. Negida A, Arnaout M, El Ashal G, Fouda S, Ghanem E, El Ghonemy S. Meta-analysis of mortality following Subthalamic and Pallidal deep brain stimulation for patients with Parkinson's disease. Poster, Association of Surgeons in Training conference, 27 Feb-1 March 2015, Glasgow.
6. Sako W, Miyazaki Y, Izumi Y, Kaji R. Which target is best for patients with Parkinson's disease? A meta-analysis of pallidal and subthalamic stimulation. J Neurol Neurosurg Psychiatry 2014; 85:982–6.
7. Sprengers M, Vonck K, Carrette E, Marson AG, Boon P. Deep brain and cortical stimulation for epilepsy. Cochrane Database Syst Rev 2014;6:CD008497.